AD_____

Award Number: DAMD17-02-1-0300

TITLE: Structural Basis for BRCA1 Function in Breast Cancer

PRINCIPAL INVESTIGATOR: John A. Ladias, M.D.

CONTRACTING ORGANIZATION: Beth Israel Deaconess Medical Center
Boston, MA 02215

REPORT DATE: April 2003

TYPE OF REPORT: Annual

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are
those of the author(s) and should not be construed as an official
Department of the Army position, policy or decision unless so
designated by other documentation.

# REPORT DOCUMENTATION PAGE

| 1. AGENCY USE ONLY (Leave blank) | 2. REPORT DATE Apr 2003 | 3. REPORT TYPE AND DATES COVERED Annual (1 Apr 2002 – 31 Mar 2003) |
|---|---|---|

**4. TITLE AND SUBTITLE**
Structural Basis for BRCA1 Function in Breast Cancer

**5. FUNDING NUMBERS**
DAMD17-02-1-0300

**6. AUTHOR(S)**
John A. Ladias, M.D.

**7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)**
Beth Israel Deaconess Medical Center
Boston, MA 02115

E-Mail: jladias@bidmc.havard.edu

**8. PERFORMING ORGANIZATION REPORT NUMBER**

**9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)**

U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

**10. SPONSORING / MONITORING AGENCY REPORT NUMBER**

**11. SUPPLEMENTARY NOTES**
Original contains color plates; All DTIC reproductions will be in black and white.

**12a. DISTRIBUTION / AVAILABILITY STATEMENT**
Approved for Public Release; Distribution Unlimited

**12b. DISTRIBUTION CODE**

**13. ABSTRACT (Maximum 200 Words)**

The Breast Cancer Susceptibility gene 1 (BRCA1) encodes an 1863-amino acid protein that plays a central role in the pathogenesis of hereditary breast cancer. The BRCA1 protein contains an N-terminal RING finger and two C-terminal BRCT domains (BRCT1 and BRCT2), which are critical for BRCA1-mediated tumor suppression and are targets for cancer-causing mutations. The BRCA1 RING interacts with the RING domain of BARD1, another protein involved in breast cancer pathogenesis, and with the C-terminal domain of BAP1 (amino acids 598-729), a ubiquitin hydrolase that enhances BRCA1-mediated cell growth suppression, whereas the BRCA1 BRCT domains interact with the C-terminal region of BACH1 (amino acids 888-1063), a helicase-like protein that contributes to the BRCA1 DNA repair function. The present project focuses on the elucidation of the structural basis of the BRCA1 RING and BRCT domain interaction with the proteins BARD1, BAP1, and BACH1, using X-ray crystallography. In the first year of the award we completed the Tasks 1a, 1b, and 1c of the original Statement of Work. For the remaining years of the award, we propose to determine the crystal structures of the BRCA1(residues 1-304)/BAP1(residues 598-729), BRCA1(residues 1-304)/BARD1(residues 25-189), and BRCA1(residues 1650-1863)/BACH1(residues 888-1063) protein complexes.

**14. SUBJECT TERMS**
Tumor suppressor gene, hereditary breast cancer, DNA repair, structural chemistry, x-ray crystallography

**15. NUMBER OF PAGES**
23

**16. PRICE CODE**

| 17. SECURITY CLASSIFICATION OF REPORT Unclassified | 18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified | 19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified | 20. LIMITATION OF ABSTRACT Unlimited |
|---|---|---|---|

# Table of Contents

## INTRODUCTION

The Breast Cancer Susceptibility gene 1 (BRCA1) encodes an 1863-amino acid protein that plays a central role in the pathogenesis of hereditary breast cancer (1-3). The BRCA1 protein contains an N-terminal RING finger and two C-terminal BRCT domains (BRCT1 and BRCT2), which are critical for BRCA1-mediated tumor suppression and are targets for cancer-causing mutations (1-4). The BRCA1 RING interacts with the RING domain of BARD1, another protein involved in breast cancer pathogenesis, and with the C-terminal domain of BAP1 (amino acids 598-729), a ubiquitin hydrolase that enhances BRCA1-mediated cell growth suppression, whereas the BRCA1 BRCT domains interact with the C-terminal region of BACH1 (amino acids 888-1063), a helicase-like protein that contributes to the BRCA1 DNA repair function (4-7). Importantly, a heterodimer consisting of the RING-encompassing regions of BRCA1 (amino acids 1-304) and BARD1 (amino acids 25-189) functions as a ubiquitin ligase targeting cellular proteins for destruction, whereas the individual BRCA1 and BARD1 domains have very low ubiquitin ligase activities (5). The present project focuses on the elucidation of the structural basis of the BRCA1 RING and BRCT domain function, using X-ray crystallography. The central role of BRCA1 in breast cancer pathogenesis provides a compelling reason for determining its three-dimensional structure, a prerequisite step for unraveling the molecular mechanisms of its function during cellular physiology and breast carcinogenesis.

## BODY

During the first year of the award we proceeded with the completion of the experiments proposed in Task 1 of the original Statement of Work. A description of our progress in these studies follows below.

### Tasks 1a, 1b, and 1c. Cloning of the human BRCA1 RING, BRCT1, and BRCT2, and human BARD1 RING domains in prokaryotic expression vectors, expression, purification and crystallization of the recombinant proteins.

The BRCA1 RING (amino acids 1-109) was produced in $E.\ coli$ cells as a fusion with MPB (maltose-binding protein) and purified (Figure 1A). The RING domain was released from MBP with thrombin cleavage, purified, and used for crystallization experiments (Figure 1B). The BRCA1 BRCT2 domain (amino acids 1757-1855) was produced in BL21(DE3) cells as an insoluble 6His-tagged protein, purified on Ni-NTA resin under denaturing conditions, refolded onto the beads, further purified (Figure 1C), and used for crystallization. Crystals of BRCT2 protein were obtained using the sitting drop vapor diffusion method (Figure 1D). Diffraction data were collected from a BRCT2 crystal and a good quality data set was obtained, 100% complete to 2.9 Å with a merging R-factor of 9.2% for all reflections (Figure 1E).

For the expression of recombinant BRCA1 RING domain (amino acids 1-304), the BARD1 RING domain (amino acids 25-189), and the BRCA1 BRCT1 domain (amino acids 1650-1736), DNA fragments coding for these domains were amplified using the polymerase chain reaction method and were cloned into pET prokaryotic vectors (Novagen) using standard methods (8). The resulting constructs were verified by DNA sequencing and were used to transform $E.\ coli$ BL21(DE3) cells. Expression of 6His-tagged recombinant proteins was induced by IPTG at $OD_{600}$ of 0.6 and the proteins were purified on Ni-NTA under denaturing conditions, refolded onto the beads, and further purified using gel filtration, as we described previously (9-14).
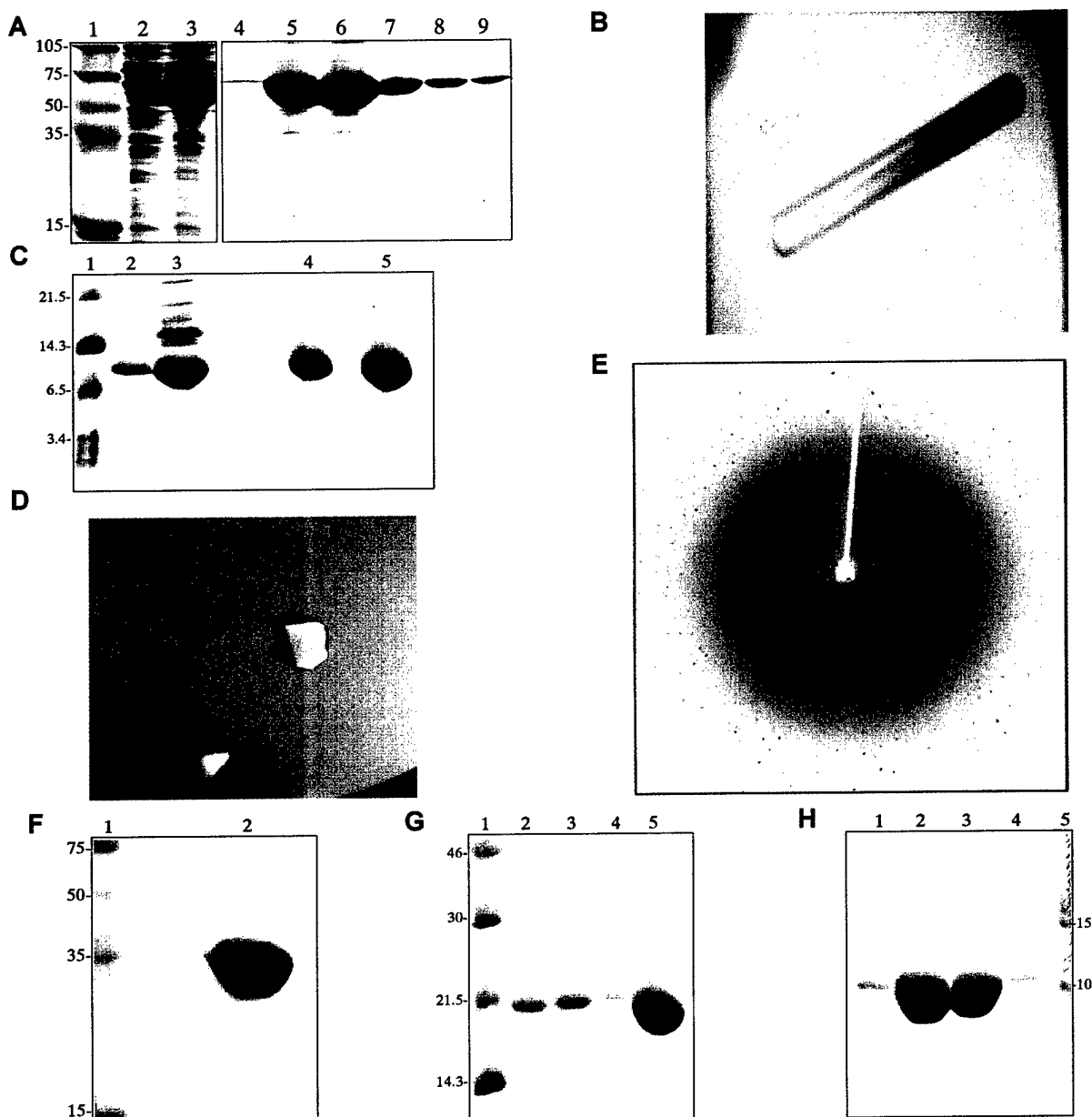
4

**Figure 1. (A)** Expression of the BRCA1 RING (amino acids 1-109). SDS-PAGE of MBP-RING fusion produced in *E. coli* cells and purified on amylose resin. *Lane 1*: protein markers; *lanes 2,3*: uninduced and induced whole-cell extracts, respectively; *lanes 4-9*: eluted fractions of MBP-RING protein. **(B)** Crystals of recombinant BRCA1 RING protein (residues 1-109) grown using the sitting drop vapor diffusion method. **(C)** Expression of BRCA1 BRCT2. SDS-PAGE of 6His-tagged BRCT2 produced in *E. coli* cells under denaturing conditions, purified and refolded onto a Ni-NTA column. *Lane 1*: protein markers; *lanes 2,3*: eluted fractions of refolded BRCT2 protein; *lanes 4,5*: refolded BRCT2 after purification by ion exchange and gel filtration chromatography. **(D)** Crystals of BRCA1 BRCT2 protein grown using the sitting drop vapor diffusion method. **(E)** Diffraction pattern of a BRCT2 crystal. The data were collected on an R-AXIS IV imaging plate detector system with 3° oscillation frames exposed for 30 min at room temperature and crystal-to-detector distance of 150 mm. **(F)** Expression of the BRCA1 RING (amino acids 1-304). SDS-PAGE of 6His-tagged protein produced in *E. coli* cells. **(G)** Expression of the BARD1 RING (amino acids 25-189). *Lanes 2-5:* various concentrations of 6His-tagged protein produced in *E. coli* cells. **(H)** Expression of BRCA1 BRCT1 (amino acids 1650-1736). *Lanes 1-4:* eluted fractions of 6His-tagged protein produced in *E. coli* cells.

Using this approach, we were able to obtain up to 5 mg of recombinant protein (>98% pure as estimated by SDS-PAGE) per liter of bacterial culture (Figure 1F-1H). The purified proteins were concentrated using Centriprep concentrators (Amicon) before they were used for crystallization experiments.

**Tasks 3a and 3b. To determine the structure of the BRCA1 RING and BRCA1 BRCT2 domains.** Although we were successful in crystallizing the BRCA1 RING protein (amino acids 1-109) and the BRCT2 domain (amino acids 1757-1855), by the time this proposal was funded (April 2002), the NMR structure of the BRCA1 RING (amino acids 1-103) complexed with the BARD1 RING (amino acids 26-122), and the crystal structure of BRCA1 BRCT1 and BRCT2 domains (amino acids 1646-1859) were published by other groups (15,16). For this reason, we did not proceed with the structural analysis of these domains as described in Tasks 3a and 3b of the original Statement of Work. Instead, we strongly believe that it is absolutely necessary to refocus and expand this project on the elucidation of the structural basis of the BRCA1 RING and BRCT domain interaction with their target proteins BARD1, BAP1, and BACH1 (1-7), using X-ray crystallography. Therefore, we need to change the original Specific Aims 1 and 3, which proposed to determine the crystal structures of the BRCA1 region (residues 1-109) and the isolated BRCA1 BRCT1 and BRCT2 domains, respectively, to the following:

**New Specific Aim 1.** To co-crystallize and determine the crystal structure of the BRCA1 RING-containing region (amino acids 1-304) complexed with the C-terminal domain of BAP1 (amino acids 598-729).

**New Specific Aim 3.** To co-crystallize and determine the crystal structure of the BRCA1 BRCT1/BRCT2-containing region (amino acids 1650-1863) complexed with the C-terminal region of BACH1 (amino acids 888-1063).

The studies proposed in the new Specific Aims 1 and 3 will reveal the interaction interfaces of the BRCA1 RING and BRCT domains with their protein targets BAP1 and BACH1, respectively. The information obtained from these experiments will elucidate the mechanisms underlying regulation of BRCA1 by BAP1 and BACH1 and will reveal the molecular changes induced by cancer-causing mutations in the BRCA1 domains that affect their interaction with BAP1 and BACH1. Notably, the crystal structures of the BRCA1 domains bound to their target proteins will provide much more information than the structures of these domains in the unbound form (e.g. the unbound BRCT domains described in ref. 16).

The **original Specific Aim 2** proposing to determine the crystal structure of the ubiquitin ligase heterodimer consisting of the BRCA1 region spanning residues **1-304** and the BARD1 region spanning residues **25-189**, remains the same for the following reasons. **First**, the protein fragments that are proposed for structural analysis in this Aim are considerably larger than the protein domains used in previous NMR studies (BRCA1 RING spanning residues **1-103** and the BARD1 RING spanning residues **26-122**, ref. 15). Therefore, they will reveal more information on the structure of these proteins than the NMR studies (15). **Second**, the BRCA1 (residues 1-304)/BARD1 (residues 25-189) heterodimer is a ubiquitin ligase (5) whereas shorter protein fragments do not have this enzymatic activity, making the structural analysis of the BRCA1(1-304)/BARD1(25-189) heterodimer absolutely essential for the understanding of its function at the

molecular level. <u>Third</u>, the proposed crystallographic studies will provide complementary information to the published solution structure of the BRCA1 RING complexed with the BARD1 RING (15).

**Study Design.** We will follow a similar research strategy to that outlined in the original proposal. Briefly, we will clone, express, purify, assemble, and co-crystallize the following protein complexes: (i) BRCA1 RING-containing region (amino acids 1-304) and C-terminal domain of BAP1 (amino acids 598-729) (Aim 1); (ii) BRCA1 RING-containing region (amino acids 1-304) and BARD1 (amino acids 25-189) (Aim 2); and (iii) BRCA1 BRCT1/BRCT2-containing region (amino acids 1650-1863) and the C-terminal region of BACH1 (amino acids 888-1063) (Aim 3). Crystallization experiments will be performed using sparse matrix crystallization methods. X-ray diffraction data sets will be collected in our macromolecular crystallography facility and at the Cornell High Energy Synchrotron Source and the crystal structures will be determined using multiwavelength anomalous diffraction methods.

As a result of the aforementioned changes in the research strategy, **we would like to request approval for the following changes in the Tasks 2 and 3 of the original Statement of Work** for the remaining award period:

*Task 2:* To clone, express, purify, and crystallize the human BRCA1, BARD1, BAP1, and BACH1 domains (months 13-24):

    a. Cloning of the BRCA1 BRCT1/BRCT2-containing region (residues 1650-1863), the BAP1 C-terminal domain (residues 598-729), and the BACH1 region (residues 888-1063) into prokaryotic expression vectors for production in *E. coli* cells (months 13-16)

    b. Expression of the encoded recombinant proteins in *E. coli* cells and purification using chromatographic methods (months 17-24)

    c. Crystallization of the proteins using sparse matrix vapor diffusion crystallization methods, seeding, limited proteolysis, and mutagenesis to improve crystallization (months 18-24)

*Task 3.* To collect diffraction data, characterize the crystals, and determine the structure of the BRCA1, BARD1, BAP1, and BACH1 domains (months 25-36):

    a. Crystal characterization, space group and unit cell dimension determination, establishment of cryogenic conditions, using our in-house X-ray facility (months 25-26)

    b. Expression, purification, and crystallization of selenomethionine-containing proteins (months 25-26)

    c. Multiwavelength anomalous diffraction data collection of selenomethionine-containing protein crystals at the CHESS synchrotron (months 27-28)

    d. Crystal structure determination of BRCA1/BAP1 complex (months 29-31)

    e. Crystal structure determination of BRCA1/BARD1 complex (months 32-34)

    f. Crystal structure determination of BRCA1/BACH1 complex (months 35-36)

## KEY RESEARCH ACCOMPLISHMENTS

1. Cloning of the BRCA1 RING domain (amino acids 1-109) into prokaryotic expression vectors *(Task 1a)*.
2. Cloning of the BRCA1 RING domain (amino acids 1-304) into prokaryotic expression vectors *(Task 1a)*.
3. Cloning of the BRCA1 BRCT1 domain (amino acids 1650-1736) into prokaryotic expression vectors *(Task 1a)*.
4. Cloning of the BRCA1 BRCT2 domain (amino acids 1757-1855) into prokaryotic expression vectors *(Task 1a)*.
5. Cloning of the BARD1 RING domain (amino acids 25-189) into prokaryotic expression vectors *(Task 1a)*.
6. Expression of the BRCA1 RING (amino acids 1-109) in *E. coli* BL21(DE3) cells and purification of the recombinant protein using chromatographic methods *(Task 1b)*.
7. Expression of the BRCA1 RING (amino acids 1-304) in *E. coli* BL21(DE3) cells and purification of the recombinant protein using chromatographic methods *(Task 1b)*.
8. Expression of the BRCA1 BRCT1 (amino acids 1650-1736) in *E. coli* BL21(DE3) cells and purification of the recombinant protein using chromatographic methods *(Task 1b)*.
9. Expression of the BRCA1 BRCT2 (amino acids 1757-1855) in *E. coli* BL21(DE3) cells and purification of the recombinant protein using chromatographic methods *(Task 1b)*.
10. Expression of the BARD1 RING (amino acids 25-189) in *E. coli* BL21(DE3) cells and purification of the recombinant protein using chromatographic methods *(Task 1b)*.
11. Crystallization of the BRCA1 RING domain (amino acids 1-109) using sparse matrix vapor diffusion crystallization methods *(Task 1c)*.
12. Crystallization of the BRCA1 BRCT2 domain (amino acids 1757-1855) using sparse matrix vapor diffusion crystallization methods *(Task 1c)*.

## REPORTABLE OUTCOMES

Publications resulted from support by this award and acknowledging this support are listed below (reprints are included in the Appendix, pages 10-23).

1. Birrane G., Chung J. and Ladias J.A.A. (2003). Novel Mode of Ligand Recognition by the Erbin PDZ Domain. *Journal of Biological Chemistry* **278**: 1399-1402.
2. Ladias J.A.A. (2003). Structural Insights into the CFTR-NHERF Interaction. *Journal of Membrane Biology* **192**: 79-88.

## CONCLUSIONS

In the first year of the award we completed the Tasks 1a, 1b, and 1c of the original Statement of Work. Because the structures of the BRCA1 RING/BARD1 RING heterodimer and the unbound BRCA1 BRCT domains were published by the time of the approval of this award, there is an absolute need to refocus and expand the studies proposed in the original Specific Aims 1 and 3 of this project. In the new Aims we propose to determine the crystal structures of the **BRCA1(1-304)/BAP1(598-729)**, **BRCA1(1-304)/BARD1(25-189)**, and **BRCA1(1650-1863)/BACH1(888-1063)** protein complexes.

# REFERENCES

1. Venkitaraman AR. (2002). Cancer susceptibility and the functions of BRCA1 and BRCA2. *Cell* **108:** 171-182
2. Scully R, Livingston DM. (2000). In search of the tumour-suppressor functions of BRCA1 and BRCA2. *Nature* **408:** 429-432
3. Welcsh PL, Owens KN, King MC. (2000). Insights into the functions of BRCA1 and BRCA2. *Trends Genet.* **16:** 69-74
4. Deng CX, Brodie SG. (2000). Roles of BRCA1 and its interacting proteins. *Bioessays* **22:** 728-737
5. Hashizume R, Fukuda M, Maeda I, Nishikawa H, Oyake D, Yabuki Y, Ogata H, Ohta T. (2001). The RING heterodimer BRCA1-BARD1 is a ubiquitin ligase inactivated by a breast cancer-derived mutation. *J. Biol. Chem.* **276:** 14537-14540
6. Jensen DE, Proctor M, Marquis ST, Gardner HP, Ha SI, Chodosh LA, Ishov AM, Tommerup N, Vissing H, Sekido Y, Minna J, Borodovsky A, Schultz DC, Wilkinson KD, Maul GG, Barlev N, Berger SL, Prendergast GC, Rauscher FJ. (1998). BAP1: a novel ubiquitin hydrolase which binds to the BRCA1 RING finger and enhances BRCA1-mediated cell growth suppression. *Oncogene* **16:** 1097-1112
7. Cantor SB, Bell DW, Ganesan S, Kass EM, Drapkin R, Grossman S, Wahrer DC, Sgroi DC, Lane WS, Haber DA, Livingston DM. (2001). BACH1, a novel helicase-like protein, interacts directly with BRCA1 and contributes to its DNA repair function. *Cell* **105:** 149-160
8. Sambrook J, Fritsch EF, Maniatis T. (1989). Molecular Cloning: A Laboratory Manual (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press)
9. Hiremath CN, Ladias JAA. (1998). Expression and purification of recombinant hRPABC25, hRPABC17, and hRPABC14.4, three essential subunits of human RNA polymerases I, II, and III. *Prot. Expr. Purif.* **13:** 198-294
10. Webster G, Leung T, Karthikeyan S, Birrane G, Ladias JAA. (2001). Crystallographic characterization of the PDZ1 domain of the $Na^+/H^+$ exchanger regulatory factor. *Acta Crystallogr.* **D57:** 714-716
11. Karthikeyan S, Leung T, Birrane G, Webster G, Ladias JAA. (2001). Crystal structure of the PDZ1 domain of the human $Na^+/H^+$ exchanger regulatory factor provides insights into the mechanism of carboxyl-terminal leucine recognition by class I PDZ domains. *J. Mol. Biol.* **308:** 963-973
12. Karthikeyan S, Leung T, Ladias JAA. (2001). Structural basis of the $Na^+/H^+$ exchanger regulatory factor PDZ1 interaction with the carboxyl-terminal region of the cystic fibrosis transmembrane conductance regulator. *J. Biol. Chem.* **276:** 19683-19686
13. Karthikeyan S, Leung T, Ladias JAA. (2002). Structural Determinants of the $Na^+/H^+$ exchanger regulatory factor interaction with the $\beta_2$ adrenergic and platelet-derived growth factor receptors. *J. Biol. Chem.* **277:** 18973-18978
14. Birrane G, Chung J, Ladias JAA. (2003). Novel Mode of Ligand Recognition by the Erbin PDZ Domain. *J. Biol. Chem.* **278:** 1399-1402
15. Brzovic PS, Rajagopal P, Hoyt DW, King MC, Klevit RE. (2001). Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. *Nat. Struct. Biol.* **8:** 833-837
16. Williams RS, Green R, Glover JN. (2001). Crystal structure of the BRCT repeat region from the breast cancer-associated protein BRCA1. *Nat. Struct. Biol.* **8:** 838-842

# *Accelerated Publication*

# Novel Mode of Ligand Recognition by the Erbin PDZ Domain*

**Gabriel Birrane, Judy Chung,
and John A. A. Ladias‡**

*From the Molecular Medicine Laboratory and
Macromolecular Crystallography Unit, Division of
Experimental Medicine, Harvard Institutes of Medicine,
Harvard Medical School, Boston, Massachusetts 02115*

**Erbin contains a class I PDZ domain that binds to the C-terminal region of the receptor tyrosine kinase ErbB2, a class II ligand. The crystal structure of the human Erbin PDZ bound to the peptide EYLGLDVPV corresponding to the C-terminal residues 1247–1255 of human ErbB2 has been determined at 1.25-Å resolution. The Erbin PDZ deviates from the canonical PDZ fold in that it contains a single α-helix. The isopropyl group of valine at position −2 of the ErbB2 peptide interacts with the Erbin Val$^{1351}$ and displaces the peptide backbone away from the α-helix, elucidating the molecular basis of class II ligand recognition by a class I PDZ domain. Strikingly, the phenolic ring of tyrosine −7 enters into a pocket formed by the extended β2-β3 loop of the Erbin PDZ. Phosphorylation of tyrosine −7 abolishes this interaction but does not affect the binding of the four C-terminal peptidic residues to PDZ, as revealed by the crystal structure of the Erbin PDZ complexed with a phosphotyrosine-containing ErbB2 peptide. Since phosphorylation of tyrosine −7 plays a critical role in ErbB2 function, the selective binding and sequestration of this residue in its unphosphorylated state by the Erbin PDZ provides a novel mechanism for regulation of the ErbB2-mediated signaling and oncogenicity.**

PDZ[1] (PSD-95/DLG/ZO-1) domains are protein interaction modules that play fundamental roles in the assembly of membrane receptors, ion channels, and other molecules into signal transduction complexes known as transducisomes (1–3). The PDZ fold comprises a six-stranded antiparallel β-barrel capped by two α-helices (1–6). PDZ domains interact with C-terminal peptides and are currently classified into two major categories based on their target sequence specificity. Class I domains bind to peptides with the consensus $X$-(S/T)-$X$-$\Phi$ ($X$ denoting any amino acid and $\Phi$ representing a hydrophobic residue), whereas class II domains recognize the motif $X$-$\Phi$-$X$-$\Phi$ (1–3). The residues at positions 0 and −2 of the peptide (position 0 referring to the C-terminal residue) play a critical role in the specificity and affinity of the interaction, whereas it is believed that amino acids upstream of the −5 position do not interact with PDZ (1–7). However, the structural determinants of ligand selectivity by PDZ domains are more complex than initially thought. For example, recent studies established an important contribution of the penultimate peptidic residue in the PDZ-ligand interaction (5, 6). Furthermore, several PDZ domains have sequence specificities that do not fall into the two classes implying the existence of more categories, whereas others bind both class I and II ligands, suggesting an intrinsic flexibility in these modules to accommodate both polar and non-polar side chains at position −2 (1–3).

Erbin was originally identified as a protein that interacts with the receptor tyrosine kinase ErbB2 (also known as HER-2 or Neu) and plays a role in its localization at the basolateral membrane of epithelial cells (8, 9). Recent studies have shown that Erbin is also highly concentrated at neuronal postsynaptic membranes and neuromuscular junctions, where it interacts with ErbB2 (10). Erbin contains a class I PDZ domain that binds with high affinity to the sequence DSWV present at the C termini of δ-catenin, ARVCF, and p0071 (11, 12). Notably, the ErbB2 sequence EYLGLDVPV that is recognized by the Erbin PDZ (8, 13), is a class II ligand, posing an interesting structural problem regarding the molecular mechanisms underlying the dual ligand specificity of this domain. The Erbin PDZ binds preferentially to the ErbB2 tail having an unphosphorylated tyrosine at position −7 (corresponding to Tyr$^{1248}$ in full-length human ErbB2), whereas phosphorylation of this residue reduces significantly the affinity of the Erbin-ErbB2 interaction (8). This preference for an unphosphorylated tyrosine is intriguing, because a PDZ interaction with the peptidic residue −7 has not been observed in previous structural studies (1–7). Importantly, phosphorylation of Tyr$^{1248}$ following ErbB2 activation is a critical event for the mitogenic signaling and oncogenicity of this receptor (14–16). Moreover, Tyr$^{1248}$ plays an important role in the basolateral localization of ErbB2 (17).

Here, we present the crystal structure of the Erbin PDZ bound to the ErbB2 C terminus. The structure reveals a novel interaction of the peptidic Tyr −7 with the extended β2-β3 loop of the Erbin PDZ. A second crystal structure of this domain bound to a phosphotyrosine-containing ErbB2 peptide shows that phosphorylation of Tyr −7 abrogates its interaction with the β2-β3 loop. These results suggest new mechanisms for regulation of the ErbB2-mediated signaling through its dynamic interaction with the Erbin PDZ.

## EXPERIMENTAL PROCEDURES

*Protein Crystallization*—A DNA fragment encoding the human Erbin PDZ domain (residues 1280–1371) was amplified from Quick-Clone cDNA (Clontech) using the polymerase chain reaction and cloned into a modified pGEX-2T vector. The Erbin PDZ was expressed in *Escherichia coli* BL21(DE3) cells as a glutathione S-transferase fusion, purified on glutathione-Sepharose 4B, released with thrombin digestion, and fur-

| | Crystal 1[a] | Crystal 2[a] | | Crystal 3[a] | |
|---|---|---|---|---|---|
| **Data collection and phasing** | | | | | |
| Data set | Native 1 | Native 2 | MAD λ1 | MAD λ2 | MAD λ3 |
| Wavelength (Å) | 1.5418 | 0.9786 | 0.9789 | 0.9686 | 0.9791 |
| Resolution (Å) | 1.88 | 1.25 | 1.4 | 1.4 | 1.4 |
| Unique reflections | 7,068 | 24,313 | 17,583 | 17,566 | 17,680 |
| Completeness (%)[b] | 95.0 (90.2) | 92.7 (91.4) | 97.8 (89.2) | 98.2 (89.2) | 98.5 (92.9) |
| $R_{sym}$ (%)[c] | 2.8 (9.6) | 2.5 (3.7) | 4.0 (7.3) | 3.7 (7.3) | 4.1 (7.5) |
| $\langle I \rangle / \langle \sigma \rangle$ | 44.7 (13.3) | 34.7 (16.8) | 48.1 (23.4) | 48.1 (23.2) | 47.6 (22.5) |
| **Refinement** | | | | | |
| Resolution range (Å) | 30.6–1.88 | 25–1.25 | | | |
| Reflections in working/test set | 6,737/330 | 23,073/1,240 | | | |
| $R_{cryst}$ (%)[d] | 16.7 | 12.8 | | | |
| $R_{free}$ (%)[e] | 21.6 | 16.4 | | | |
| Allowed/additional/generous (%) regions in Ramachandran plot | 88.2/9.2/2.6 | 90.8/6.6/2.6 | | | |

[a] Crystals 1 and 2 refer to crystals of the wild-type Erbin PDZ bound to the phosphorylated and unphosphorylated ErbB2 peptides, respectively. Crystal 3 refers to the SeMet-substituted PDZ(V1366M) bound to the unphosphorylated peptide.
[b] Numbers in parentheses refer to the highest resolution shell (1.93–1.88 Å for Native 1, 1.29–1.25 Å for Native 2, and 1.45–1.4 Å for MAD).
[c] $R_{sym} = \Sigma |(I - \langle I \rangle)| / \Sigma(I)$, where $I$ is the observed integrated intensity, $\langle I \rangle$ is the average integrated intensity obtained from multiple measurements, and the summation is over all observed reflections.
[d] $R_{cryst} = \Sigma \|F_{obs}| - k|F_{calc}\| / \Sigma|F_{obs}|$, where $F_{obs}$ and $F_{calc}$ are the observed and calculated structure factors, respectively.
[e] $R_{free}$ is calculated as $R_{cryst}$ using 5% of the reflection data chosen randomly and omitted from the refinement calculations.

ther purified by gel filtration (5). The Erbin PDZ protein (19 mg/ml in 500 mM NaCl, 50 mM Tris-HCl, pH 8.3) was mixed with the synthetic peptide EYLGLDVPV at an equimolar ratio and crystallized in 12–15% polyethylene glycol 4000, 10% glycerol, 100 mM ammonium acetate, 100 mM sodium acetate, pH 4.6, at 20 °C, using the sitting drop vapor diffusion method. Crystals were cryoprotected in mother liquor containing 30% glycerol and flash-frozen in a liquid nitrogen stream. The mutation V1366M was introduced in the Erbin PDZ using the polymerase chain reaction, and the resulting protein was expressed in B834(DE3)pLysS cells grown in minimal medium supplemented with 40 mg/l selenomethionine (SeMet). The SeMet-protein was purified and co-crystallized with the ErbB2 peptide under similar conditions. Multiwavelength anomalous dispersion (MAD) data sets of the SeMet-substituted PDZ(V1366M)-peptide crystals were collected at 100 K using synchrotron radiation at the Cornell High Energy Synchrotron Source (F2 station), Ithaca, NY. High resolution data of isomorphous crystals of the wild-type Erbin PDZ-peptide complex were also collected at the F2 station. The crystals belong to space group P2₁ with unit cell dimensions $a = 26.6$ Å, $b = 57.4$ Å, $c = 30.4$ Å, $\beta = 100.6°$. Crystals of the Erbin PDZ bound to the peptide EpYLGLDVPV (pY denoting phosphotyrosine) were obtained under similar conditions and were analyzed at 100 K using CuKα radiation. The crystals belong to space group P2₁ with $a = 26.5$ Å, $b = 57.0$ Å, $c = 30.9$ Å, $\beta = 99.2°$. Data were processed using DENZO and SCALEPACK (18) (Table I).

*Structure Determination and Refinement*—The crystal structure of the SeMet-substituted PDZ(V1366M)-peptide complex was determined using SOLVE/RESOLVE (19). The obtained phases were used to solve the structure of the wild-type Erbin PDZ-peptide complex at 1.25-Å resolution. Phase extension and automated model building were performed using wARP (20), in combination with manual intervention using O (21). Initial isotropic refinement was performed using REFMAC (22), followed by several rounds of anisotropic refinement with SHELXL-97 (23). The structure of the Erbin PDZ-phosphopeptide complex was determined by molecular replacement with AMoRe (24) using the Erbin PDZ as the search model. The crystallized PDZ domain includes the vector-derived residues GSM at its N terminus. In the 1.25-Å structure the side chains of PDZ residues Glu[1280], Ser[1294], Ser[1325], His[1347], Gln[1349], and Ile[1365] are modeled in two conformations. The main conformation of His[1347] (occupancy 0.7) has excellent electron density and is used to describe the present structure, whereas the electron density for the minor conformation is of poor quality.

*Isothermal Titration Calorimetry*—Binding constants of the Erbin PDZ to the ErbB2 peptides were measured using a VP-ITC microcalorimeter (MicroCal, LLC). Briefly, a 0.896 mM solution of the native and a 0.830 mM solution of the phosphotyrosine-containing ErbB2 peptide were titrated into a 0.0389 mM solution of Erbin PDZ protein in 25 mM Tris-HCl, pH 8.3, at 25 °C. Titration curves were analyzed using the program ORIGIN 5.0 (OriginLab).

## RESULTS AND DISCUSSION

*Structural Basis for Class II Ligand Recognition by the Erbin PDZ*—The crystal structure of the Erbin PDZ bound to the

ErbB2 peptide EYLGLDVPV was determined using MAD phasing and was refined anisotropically to 1.25-Å resolution. The Erbin PDZ lacks the short α-helix that is present between the β3 and β4 strands in PDZs with known structure (1–7), due to the shorter length of the Erbin β3-β4 loop (Fig. 1, *A* and *B*). The significance of this deviation from the canonical PDZ fold is unclear because this α-helix has no known function (1–3) and its inconsequential absence from the Erbin PDZ argues against a structural role in the folding of this module.

The ErbB2 peptide inserts into the Erbin PDZ ligand-binding groove antiparallel to the β2 strand, extending and twisting the β-sheet of PDZ (Fig. 1, *B* and *C*). The isopropyl and carboxylate groups of Val 0 enter into the carboxylate-binding pocket (designated here as P₁), where they are stabilized through hydrophobic interactions and hydrogen bonds with PDZ residues (Fig. 1*D*), similar to those described for other class I PDZ-ligand complexes (1–6). Remarkably, the isopropyl group of Val −2 makes hydrophobic contacts with Val[1351], which appear to cause a displacement of the peptide backbone away from the α-helix (Fig. 1*B*), providing an explanation for the ability of Erbin PDZ to recognize a class II ligand. The peptide is further stabilized at this position through an interaction of Asp −3 with Thr[1316] (Fig. 1*D*), whereas Leu −4, Gly −5, and Leu −6 do not bind to PDZ. Interestingly, the imidazole ring of the conserved His[1347], which is the hallmark of class I PDZ domains and plays a critical role in the selection of the residue −2, points away from Val −2, where it hydrogen bonds with the carbonyl oxygen of Gly[1299] (Fig. 1*B*).

*The β2-β3 Loop of Erbin PDZ Interacts with Tyr −7 of the ErbB2 Ligand*—Strikingly, the phenolic ring of Tyr −7 folds back in a direction parallel to the peptidic backbone and enters a pocket, designated P₂, which is formed by Ser[1296] in the β2 strand and Gly[1303], Asn[1304], and Pro[1305] in the β2-β3 loop (Fig. 1, *B* and *C*). This represents the first structural evidence for a direct interaction of the PDZ domain with the peptidic residue −7. The β2-β3 loop of Erbin PDZ is considerably longer than that of PDZs with known structure (Fig. 1*A*) and contains five glycine and two proline residues that create a bent platform against which Tyr −7 is stacked. The phenolic ring of Tyr −7 is stabilized mainly by hydrophobic interactions and is well ordered, as indicated by the high quality electron density map (Fig. 1*E*). The hydroxyl group of Tyr −7 hydrogen bonds through two ordered water molecules with Asp −3 (Fig. 1*D*).

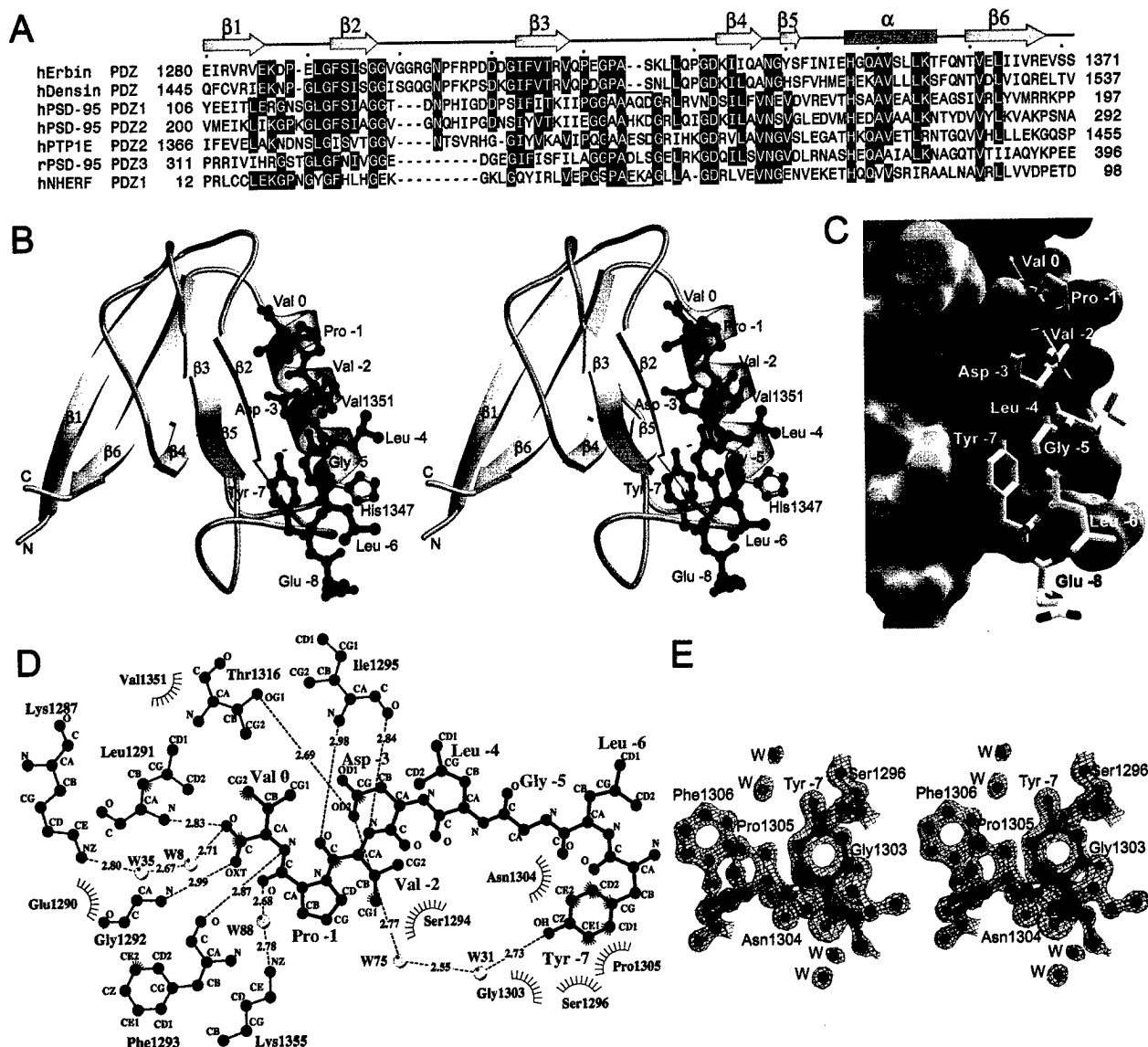*Phosphorylation of Tyr −7 Abolishes Binding to the P₂ Pock-*

FIG. 1. **Structure of the Erbin PDZ bound to the unphosphorylated ErbB2 peptide.** *A*, sequence comparison of selected class I PDZ domains. Identical residues in four or more domains are shown as *white* letters on *blue* background. *Hyphens* represent gaps inserted for optimum alignment. The secondary structure of the Erbin PDZ is indicated at the *top*. Residues forming a short α-helix in PDZs with known structures are enclosed in a *red box*. *B*, stereo view of the Erbin PDZ bound to the peptide EYLGLDVPV. The figure was made using BOBSCRIPT (30) and POV-Ray (www.povray.org). *C*, surface topology of the Erbin PDZ bound to the ErbB2 peptide. The figure was made using GRASP (31). *D*, two-dimensional representation of the interactions between Erbin PDZ residues (*orange*) and the peptide (*purple*). Water molecules (*W*) are shown as *cyan spheres*, hydrogen bonds as *dashed lines*, and hydrophobic interactions as *arcs* with *radial spokes*. The figure was made using LIGPLOT (32). *E*, stereo view of a weighted $2F_{obs} - F_{calc}$ electron density map at the $P_2$ pocket calculated at 1.25 Å and contoured at 2.5 σ.

*et*—Because phosphorylation of Tyr[1248] plays a critical role in ErbB2 signaling (14–16), we also determined the crystal structure of the Erbin PDZ bound to the peptide EpYLGLDVPV. No electron density is observed for the peptidic residues −5 to −8 and the $P_2$ pocket is empty (Fig. 2A). In contrast, Val 0, Pro −1, Val −2, Asp −3, and Leu −4 are well ordered inside the ligand-binding groove (Fig. 2A). The integrity of the peptide in the crystallized complex was verified by mass spectroscopic analysis (data not shown), indicating that the invisible portion of the peptide is disordered and faces toward the solution. Isothermal titration calorimetry experiments showed that the native ErbB2 peptide binds to the Erbin PDZ with a $K_d$ of ~50 μM, whereas the phosphotyrosine-containing peptide binds to PDZ with a $K_d$ of ~128 μM. The ~2.5-fold reduction in the affinity of Erbin PDZ for the phosphorylated ErbB2 peptide is attributed to the loss of the hydrophobic interactions and hydrogen bonds stabilizing the phenolic ring of Tyr −7 inside the $P_2$ pocket.

Superposition of the Erbin PDZ structures with the PSD-95 PDZ3 (4) reveals that Val 0, Pro −1, Val −2, and Asp −3 are superposed extremely well in both Erbin complexes, whereas the ErbB2 backbone is displaced away from the α-helix as compared with PSD-95 PDZ3 (Fig. 2B). These results indicate that the displacement of the ErbB2 peptide is due to the Val −2 interaction with Val[1351] rather than the Tyr −7 binding to $P_2$. Only small differences are observed in the backbone positions of the Erbin β2-β3 loops (overall root-mean-square deviation 0.26 Å for residues 1299–1311), indicating that the $P_2$ site is preformed and does not undergo major conformational changes upon Tyr −7 binding. By contrast, the β2-β3 loops of the Erbin PDZ and PSD-95 PDZ3 occupy completely different positions and are not superimposable.

*Structural and Functional Implications*—The property of the newly discovered pocket $P_2$ to discriminate between the phosphorylation states of Tyr −7 indicates that it may play a regulatory role in ErbB2 signaling and suggests an attractive
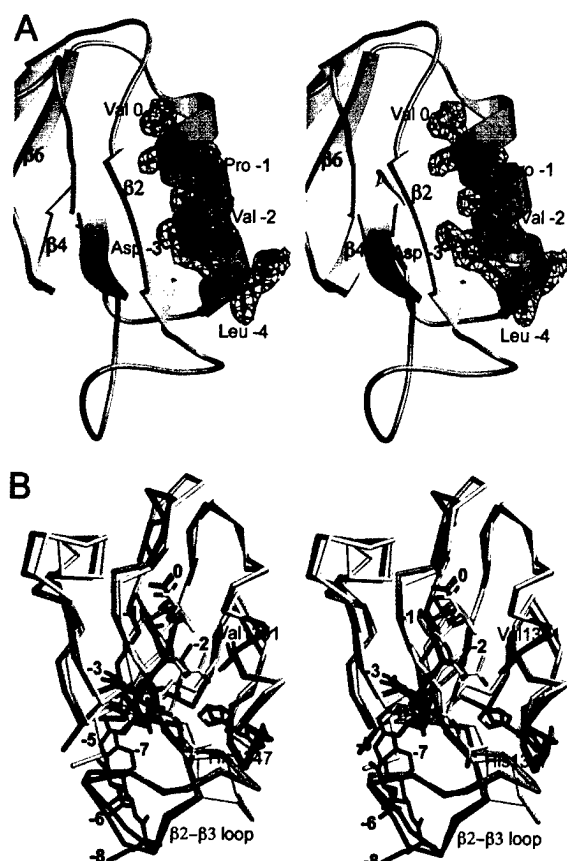
**FIG. 2. Structure of the Erbin PDZ bound to the phosphorylated ErbB2 peptide.** *A*, stereo view of the Erbin PDZ bound to the peptide EpYLGLDVPV. A weighted $2F_{obs} - F_{calc}$ electron density map calculated at 1.88-Å resolution and contoured at 1.0 $\sigma$ is superimposed on the ErbB2 peptide. *B*, superposition of the Cα backbone traces of Erbin PDZ-peptide (*pink*), Erbin PDZ-phosphopeptide (*blue*), and PSD-95 PDZ3-peptide (*yellow*) (Protein Data Bank code 1BE9). Side chains of the peptidic residues, Erbin His[1347] and Val[1351], and PSD-95 His[372] are shown as *stick* models.

model for this regulation. Conceivably, during the inactive state of ErbB2, Tyr −7 is buried inside $P_2$ and is inaccessible for phosphorylation and interaction with other proteins. Activation of ErbB2 triggers the release of Tyr −7 from $P_2$, possibly through conformational changes induced in Erbin and/or the cytoplasmic domain of ErbB2. Notably, Erbin becomes phosphorylated by ErbB2 following receptor activation (8), raising the intriguing possibility that this may represent a step preceding the dissociation of Tyr −7. Subsequently, the released tyrosine is primed for phosphorylation and interaction with phosphotyrosine-binding domains (*e.g.* PTB or SH2) of downstream signaling proteins (14, 15). Following signal transduction, dephosphorylation of Tyr −7 restores its original position inside $P_2$. Importantly, in contrast to the regulatory site $P_2$ that oscillates between bound and unbound states, $P_1$ interacts constitutively with the last four residues of ErbB2 securing the continuous participation of Erbin and ErbB2 in the same macromolecular complex at the basolateral membrane throughout the activation-inactivation cycles of the receptor. This model also allows for simultaneous binding of the Erbin PDZ and either PTB or SH2 domains to the phosphorylated ErbB2 C-terminal region, because these modules have non-overlapping recognition motifs.

Do other PDZ domains have a $P_2$ pocket? In contrast to the short β2-β3 loops of PSD-95 PDZ3 and NHERF PDZ1 (Fig. 1A)

that have not been shown to interact with peptidic residues (4–6), the extended β2-β3 loops of the PSD-95 PDZ1, PSD-95 PDZ2, and PTP1E PDZ2 domains are involved in ligand interactions (7, 25–28). Importantly, alternative spliced isoforms of PTP1E PDZ2 with different β2-β3 loop lengths have entirely different binding affinities for the C-terminal region of the tumor suppressor protein APC (29), providing further evidence for an important role of $P_2$ in PDZ-ligand interactions. These observations, taken together with the present structures of Erbin PDZ, demonstrate that the $P_2$ site is a hitherto unrecognized important structural element with possible regulatory function, at least for a subset of PDZ domains. Moreover, the emerging complexity of PDZ selectivity mechanisms points to the need for new PDZ classification schemes that will take into consideration the β2-β3 loop length, the specificity of the $P_2$-ligand interaction, and the structural determinants underlying the dual ligand specificity of these versatile protein modules.

REFERENCES

1. Sheng, M., and Sala, C. (2001) *Annu. Rev. Neurosci.* **24**, 1–29
2. Harris, B. Z., and Lim, W. A. (2001) *J. Cell Sci.* **114**, 3219–3231
3. Hung, A. Y., and Sheng, M. (2002) *J. Biol. Chem.* **277**, 5699–5702
4. Doyle, D. A., Lee, A., Lewis, J., Kim, E., Sheng, M., and MacKinnon, R. (1996) *Cell* **85**, 1067–1076
5. Karthikeyan, S., Leung, T., and Ladias, J. A. A. (2001) *J. Biol. Chem.* **276**, 19683–19686
6. Karthikeyan, S., Leung, T., and Ladias, J. A. A. (2002) *J. Biol. Chem.* **277**, 18973–18978
7. Kozlov, G., Banville, D., Gehring, K., and Ekiel, I. (2002) *J. Mol. Biol.* **320**, 813–820
8. Borg, J. P., Marchetto, S., Le Bivic, A., Ollendorff, V., Jaulin-Bastard, F., Saito, H., Fournier, E., Adelaide, J., Margolis, B., and Birnbaum, D. (2000) *Nat. Cell Biol.* **2**, 407–414
9. Bryant, P. J., and Huwe, A. (2000) *Nat. Cell Biol.* **2**, E141–E143
10. Huang, Y. Z., Wang, Q., Xiong, W. C., and Mei, L. (2001) *J. Biol. Chem.* **276**, 19318–19326
11. Laura, R. P., Witt, A. S., Held, H. A., Gerstner, R., Deshayes, K., Koehler, M. F., Kosik, K. S., Sidhu, S. S., and Lasky, L. A. (2002) *J. Biol. Chem.* **277**, 12906–12914
12. Jaulin-Bastard, F., Arsanto, J. P., Le Bivic, A., Navarro, C., Vely, F., Saito, H., Marchetto, S., Hatzfeld, M., Santoni, M. J., Birnbaum, D., and Borg, J. P. (2002) *J. Biol. Chem.* **277**, 2869–2875
13. Jaulin-Bastard, F., Saito, H., Le Bivic, A., Ollendorff, V., Marchetto, S., Birnbaum, D., and Borg, J. P. (2001) *J. Biol. Chem.* **276**, 15256–15263
14. Rubin, I., and Yarden, Y. (2001) *Ann. Oncol.* **12**, S3–S8
15. Dankort, D., Jeyabalan, N., Jones, N., Dumont, D. J., and Muller, W. J. (2001) *J. Biol. Chem.* **276**, 38921–38928
16. Ben-Levy, R., Paterson, H. F., Marshall, C. J., and Yarden, Y. (1994) *EMBO J.* **13**, 3302–3311
17. Dillon, C., Creer, A., Kerr, K., Kumin, A., and Dickson, C. (2002) *Mol. Cell. Biol.* **22**, 6553–6563
18. Otwinowski, Z., and Minor, W. (1997) *Methods Enzymol.* **276**, 307–326
19. Terwilliger, T. C., and Berendzen, J. (1999) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **55**, 849–861
20. Lamzin, V. S., and Wilson, K. S. (1997) *Methods Enzymol.* **277**, 269–305
21. Jones, T. A., Zou, J. Y., Cowan, S., and Kjeldgaard, M. (1991) *Acta Crystallogr. Sect. A* **47**, 110–119
22. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **53**, 240–255
23. Sheldrick, G. M., and Schneider, T. R. (1997) *Methods Enzymol.* **277**, 319–343
24. Navaza, J. (1994) *Acta Crystallogr. Sect. D Biol. Crystallogr.* **50**, 157–163
25. Piserchio, A., Pellegrini, M., Mehta, S., Blackman, S. M., Garcia, E. P., Marshall, J., and Mierke, D. F. (2002) *J. Biol. Chem.* **277**, 6967–6973
26. Tochio, H., Hung, F., Li, M., Bredt, D. S., and Zhang, M. (2000) *J. Mol. Biol.* **295**, 225–237
27. Imamura, F., Maeda, S., Doi, T., and Fujiyoshi, Y. (2002) *J. Biol. Chem.* **277**, 3640–3646
28. Walma, T., Spronk, C. A., Tessari, M., Aelen, J., Schepens, J., Hendriks, W., and Vuister, G. W. (2002) *J. Mol. Biol.* **316**, 1101–1110
29. Erdmann, K. S., Kuhlmann, J., Lessmann, V., Herrmann, L., Eulenburg, V., Muller, O., and Heumann, R. (2000) *Oncogene* **19**, 3894–3901
30. Esnouf, R. M. (1997) *J. Mol. Graph. Model.* **15**, 132–134
31. Nicholls, A., Sharp, K. A., and Honig, B. (1991) *Proteins* **11**, 281–296
32. Wallace, A. C., Laskowski, R. A., and Thornton, J. M. (1995) *Protein Eng.* **8**, 127–134

## Topical Review

# Structural Insights into the CFTR-NHERF Interaction

## J.A.A. Ladias

Molecular Medicine Laboratory and Macromolecular Crystallography Unit, Division of Experimental Medicine, Harvard Institutes of Medicine, Harvard, Medical School, 4 Blackfan Circle, Boston Massachusetts 02115, USA

## Introduction

The cystic fibrosis transmembrane conductance regulator (CFTR) is a low-conductance chloride channel located at the apical membrane of epithelial cells where it mediates cAMP-dependent chloride secretion (reviewed in [2, 23, 71]). Abnormal CFTR function is associated with the pathogenesis of human diseases, including cystic fibrosis, secretory diarrhea, and pancreatitis [2, 20]. Cystic fibrosis is caused by mutations in the CFTR gene that decrease the cell surface expression and/or activity of the CFTR protein and is the most common lethal genetic disease in Caucasians. By contrast, secretory diarrhea is caused by overstimulation of CFTR in intestinal epithelial cells by bacterial enterotoxins and is the second largest cause of infant mortality in the developing world. Therefore, the elucidation of the molecular mechanisms underlying CFTR regulation will not only provide a deep understanding of transepithelial electrolyte transport but it will also facilitate the development of novel clinical treatments for CFTR-associated diseases.

CFTR belongs to the ABC (ATP-Binding Cassette) membrane transporter superfamily (subfamily C, member ABCC7) (reviewed in [7, 13, 30, 33, 66]). ABC transporters use the chemical energy of ATP to move diverse sets of solutes across the membrane, including amino acids, peptides, large proteins, lipids, sugars, pigments, and anions. These transporters represent the largest gene superfamily in many sequenced microbial genomes and share a common architectural organization comprising two cytoplasmic ABC domains with ATPase activity and two membrane-spanning domains each consisting of six or more transmembrane segments (TM). These four do-

mains may be expressed as separate polypeptides or half-transporters, such as the putative lipid A transporter MsbA from Escherichia coli [8], or fused together in a single polypeptide, as in the CFTR protein. The energy of ATP binding and hydrolysis is used to select and transport the substrates through the lipid bilayer by unknown mechanisms. The current draft of the human genome sequence contains 48 ABC genes, 16 of which encode transporters with known function, including the multidrug resistance protein, the transporter for antigen presentation, and the sulfonylurea receptor (a compilation of the known human ABC transporters can be found at *http://www.nutrigene.4t.com/humanabc.htm*). Because of the central role of ABC transporters in bacterial virulence and serious human disorders, including cystic fibrosis, multidrug resistance, hypercholesterolemia, Stargardt disease, and adrenoleukodystrophy [7, 13, 33, 66], a thorough understanding of the molecular mechanisms underlying their function is clinically important.

Here I will discuss briefly recent developments in the structural analysis of ABC transporters that have provided mechanistic insights into the CFTR regulation, and will examine in more depth the structural determinants of the interaction between the $Na^+/H^+$ exchanger regulatory factor (NHERF) and the C-terminal tail of CFTR, as revealed by recent crystallographic studies. Because the structural analysis of CFTR is an underdeveloped field, I will also discuss future research directions that are urgently needed to elucidate the molecular basis of CFTR function and its regulation by other proteins.

## Molecular Architecture of CFTR and ABC Transporters

CFTR is a unique member of the ABC superfamily in that it is an ATP-regulated chloride channel and not a transporter. The 1480-amino-acid CFTR protein consists of two homologous halves, each containing

six TMs connected with extracellular and cytoplasmic "loops" and a nucleotide-binding domain (NBD) (another name for the ABC used in the CFTR literature). The two halves are linked by a cytoplasmic regulatory domain ($R$) that contains many consensus sites for phosphorylation by protein kinase A (PKA), C (PKC), and cGMP-dependent protein kinase II [2, 23, 71]. Remarkably, when the two CFTR halves are expressed as separate proteins in the same cells, they assemble into functional channels, indicating that covalent linkage of the two halves is not required for channel assembly and function [2, 71].

To date, high-resolution structures of the entire CFTR channel or its domains do not exist. This lack of structural information has hampered the elucidation of mechanisms underlying CFTR function at the molecular level since the cloning of the *CFTR* gene in 1989 [63]. In the absence of a three-dimensional atomic model of CFTR, the proposed channel topology and interdomain relationships are inferred from structure-function analyses. The CFTR channel is thought to have a large extracellular vestibule that extends into the membrane, whereas the selectivity filter is located in the cytoplasmic part of the channel where the pore becomes narrow. Although the location of the gate that regulates ion conduction through the channel is currently unknown, it is well established that the channel gating is controlled by conformational changes in the cytoplasmic domains [2, 71].

Notwithstanding the lack of information on the CFTR structure, a major breakthrough in the three-dimensional organization of ABC transporters was recently achieved by the crystal structure determination of the MsbA from *E. coli* at 4.5 Å resolution [8]. The MsbA structure provided the first model for the transmembrane domain and cytoplasmic loops of a complete ABC transporter, as well as the topological relationship of these regions to the ABC domains. This structure established the mode of interaction of the α-helical TM segments with each other and revealed that the cytoplasmic "loops" are actually extensions of the TM α-helices into the cytoplasm. Surprisingly, the MsbA ABC domains are positioned remotely from each other and thus are unable to associate upon ATP binding. Although this arrangement could reflect an active conformation of the transporter, it has raised the possibility that the MsbA structure does not represent a physiologic dimer [31, 75]. In this context, it is important to note that the molecular envelope of the MsbA crystal structure differs substantially from those observed for the multidrug resistance protein 1 and the transporter for antigen presentation obtained by electron microscopy of single particles [64, 77]. Despite the limitations of the MsbA structure to represent transporters that translocate hydrophilic substrates [8], it nonetheless provides a starting structural framework that will guide future experiments toward a better understanding of the mechanochemistry of ABC transporters.

The 4.5 Å electron density map of MsbA did not reveal the ABC structure at high resolution. However, insights into the ABC fold were provided by crystal structures of several ABC domains, including the histidine periplasmic permease (HisP) from *Salmonella typhimurium* [36], the trehalose/maltose transporter (MalK) from *Thermococcus litoralis* [14], the human transporter for antigen presentation TAP1 [24], as well as the MJ1267 [37] and MJ0796 [82] ABCs from *Methanococcus jannaschii*. In addition, the crystal structure of the ATPase domain of the DNA repair enzyme Rad50 from *Pyrococcus furiosus* revealed a homodimer induced upon binding to ATP [34]. Rad50 is distantly related to ABC transporters but its ATPase domain is structurally similar to those of ABC members, primarily at the ATP-binding site. Interestingly, a recent crystal structure of the vitamin $B_{12}$ transporter BtuCD from *E. coli* at 3.2 Å resolution [45] revealed that its ABC domains contact each other in an arrangement similar to the Rad50 ATPase dimer but different from that observed for the MsbA protein. These studies established the conserved tertiary structure of the NBD/ABC fold, which comprises a core α/β subdomain containing the consensus nucleotide-binding motifs Walker A ($GX_4GKS/T$) (X denoting any amino acid) and Walker B ($RX_{6-8}\Phi_4D$) ($\Phi$ representing a hydrophobic residue), an antiparallel β subdomain that interacts with the base of the nucleotide, and an α subdomain that contains the ABC transporter signature sequence LSGGQ. The Walker A motif follows a β-strand and it forms a loop (P-loop) that wraps around the α- and β-phosphates of the nucleotide, followed by an α-helix. The Walker B motif forms a β-strand followed usually by a glutamate and it may help coordinate the $Mg^{2+}$ ion possibly through a water molecule [34] or it may polarize the attacking water molecule [36]. The function of the signature motif LSGGQ has not been determined unambiguously and it may act as a γ-phosphate sensor in the opposing molecule of the ABC dimer [34] and/or may signal to the membrane-spanning domains [33]. Although these structures provided important insights into the molecular basis of ABC transporter function, more studies are needed to define the ATP-dependent conformational changes of the ABC domains that underlie the functional cycles of these transporters and the gating of the CFTR channel. For example, there is controversy concerning the structure of the ABC dimer, which is thought to be a conserved feature of ABC transporters. The aforementioned crystal structures of isolated ABCs revealed several potential, albeit mutually exclusive, dimeric arrangements of these domains and failed to resolve unequivocally their oligomeric organization.

## Regulation of CFTR Gating

The CFTR channel gating is thought to be controlled by three distinct processes: i) phosphorylation of the R domain; ii) binding and hydrolysis of ATP by the NBDs; and iii) interactions of CFTR domains among themselves and with other proteins (reviewed in [2, 23, 42, 71]). Phosphorylation of the R domain is a prerequisite for channel activation, and the channel open probability is directly related to the extent of phosphorylation. Deactivation of CFTR is brought about by protein phosphatases, including PP2A and PP2C [46]. However, phosphorylation is not sufficient for CFTR activation. A second mechanism for the control of channel gating involves the binding and hydrolysis of ATP by the two NBDs. Numerous studies have provided evidence that the CFTR NBDs play different but cooperative roles in controlling channel gating [2, 23, 71]. These domains share limited overall sequence similarity (less than 30% amino-acid identity) and exhibit sequence variations even in the Walker and signature motifs. For example, a conserved glutamate at the end of the Walker B motif that activates the hydrolytic water for attack on the $\gamma$-phosphate of ATP is replaced by a serine in the CFTR NBD1, suggesting that NBD1 may not hydrolyze ATP efficiently. Indeed, it was recently shown that ATP binds stably and dissociates slowly from the CFTR NBD1, while it is rapidly hydrolyzed by the NBD2 [3], demonstrating the non-equivalency of these NBDs. In addition, the signature motif LSHGH of CFTR NBD2 deviates from the consensus LSGGQ. This asymmetry of sequence conservation in the ATP-binding and active sites of the CFTR NBDs may reflect the different roles of these domains in channel gating, as have been demonstrated by many biochemical studies [2, 23, 71]. Elucidation of the structural basis of the NBD1-NBD2 interactions, their functional asymmetry and cyclic conformational changes that control CFTR gating awaits crystallographic analysis of these domains in the apo form, as well as complexed with nucleotides.

In addition to phosphorylation and ATP hydrolysis, recent studies have revealed a third mechanism of CFTR regulation operating through interactions of the CFTR domains among themselves and with other proteins. Specifically, it has been demonstrated that the N-terminal domain (NTD) of CFTR interacts directly with the R domain and functions as a positive regulator of the channel activity [42, 55]. At least part of this regulatory function has been mapped to a cluster of acidic residues in the CFTR NTD, whose sequential removal results in a graded inhibition of CFTR activity [21, 55]. Furthermore, syntaxin 1A, a membrane protein that plays a central role in neurotransmitter release, binds to the CFTR NTD and inhibits channel activity [42, 56–58]. In fact, it appears that the CFTR channel is regulated through binding of its NTD to a protein complex composed of syntaxin 1A and the synaptosome-associated protein SNAP-23 [12]. Remarkably, two recent studies have shown that the channel gating is also modulated through association of the CFTR C-terminal domain (CTD) with NHERF [61] and the CFTR-associated protein-70 (CAP70) [80] also known as PDZK1 [43]. These proteins interact with the CFTR C-terminal tail through a pair of PDZ (PSD-95/Discs-large/ZO-1) domains and it is thought that they activate the channel probably by inducing and/or stabilizing its dimerization [61, 80]. However, the oligomeric state of the functional CFTR channel is currently unknown. Early studies failed to co-immunoprecipitate biochemically different CFTR proteins expressed in the same cells, suggesting that CFTR is a monomer [48]. By contrast, functional analyses of coexpressed CFTR molecules with distinct properties, as well as electron micrographs of CFTR particles led to the conclusion that CFTR is a homodimer [18, 83]. The monomer hypothesis was reinforced by recent biochemical and electrophysiological experiments [10] although these studies could not exclude the possibility that CFTR channels are transiently tethered together by other proteins to form larger macromolecular complexes. Therefore, this issue remains controversial and definitive proof of the CFTR quaternary structure, as well as elucidation of the molecular mechanisms underlying CFTR activation by NHERF and CAP70 await future structural and functional studies.

## Role of PDZ-containing Proteins in CFTR Apical Localization and Function

PDZ domains are protein interaction modules that play fundamental roles in the assembly of membrane receptors, ion channels, and other signaling molecules into specific signal transduction complexes [19, 28, 35, 68]. Such macromolecular complexes organized by PDZ-containing proteins have been termed transducisomes [19, 76] and are thought to increase the speed and specificity of signal transmission from membrane receptors to physically coupled downstream signaling molecules. The PDZ fold comprises a six-stranded antiparallel $\beta$-barrel capped by two $\alpha$-helices. C-terminal peptides interact with PDZ domains by a $\beta$-sheet augmentation process, in which the peptide forms an additional antiparallel $\beta$-strand in the PDZ $\beta$-sheet. Early studies categorized PDZ domains based on their target sequence specificity into class I domains that bind to peptides with the consensus X-(S/T)-X-$\Phi$ and class II domains that recognize the motif X-$\Phi$-X-$\Phi$ [28, 35, 68, 73]. Those studies pointed to the importance of peptide residues at positions 0 and $-2$ for the specificity and affinity (position 0 referring to the C-terminal residue),

whereas the residue −1 was thought to play no role in the interaction. This conclusion was corroborated by initial structural analyses, which showed that the side chain of the penultimate peptidic residue was facing towards the solution and did not bind to PDZ [15]. However, it became clear from subsequent studies that the structural determinants of the PDZ-ligand interaction are more complex than initially thought. For example, several PDZ domains have specificities that do not fall into these classes, implying the existence of more PDZ categories [4, 28], whereas others bind both class I and II ligands, suggesting an intrinsic flexibility in these modules to accommodate both polar and nonpolar side chains at position −2 [4, 35]. Furthermore, certain PDZ domains can also interact with internal protein sequences that adopt a β-hairpin structure [32]. Although the structural basis for ligand selection by PDZ domains is not well understood and reclassification of PDZs based on structural and affinity studies seems likely in the future, in this discussion I will use the current classification scheme of these domains.

NHERF is a cytoplasmic protein that was originally cloned as an essential cofactor for the PKA-mediated inhibition of the $Na^+/H^+$ exchanger 3 from the renal brush border [54, 81]. NHERF, also known as EBP50 (ezrin/radixin/moesin-binding phospho-protein-50) [62], contains two tandem class I PDZ domains that interact differentially with numerous target proteins and a C-terminal ezrin/radixin/moesin-binding module that associates with the cortical actin cytoskeleton. The PDZ domains of NHERF and its related protein NHERF2 (also known as E3KARP) promote homo- and heterotypic protein-protein interactions, thereby orchestrating the clustering of ion channels and membrane receptors into transducisomes at the apical plasma membrane [70, 78]. The N-terminal PDZ domain of NHERF (designated PDZ1) spans residues 11–94 and binds to the C-terminal tails of several membrane receptors and ion channels, including the sequences QDTRL, NDSLL, and EDSFL, of CFTR, $β_2$ adrenergic receptor ($β_2AR$), and platelet-derived growth factor receptor (PDGFR), respectively [26, 27, 49]. The NHERF PDZ2 domain (residues 150–235) recognizes different sequence motifs than PDZ1 [79] and very few PDZ2 targets have been identified so far, including the c-Yes-associated protein YAP-65 [51] and the chloride channel ClC-3B [60].

The PDZ-binding motif DTRL of the CFTR C-terminal tail is essential for anchoring this chloride channel to the apical membrane because its deletion results in mislocalization of CFTR in airway and kidney epithelial cells [52, 53, 72]. The importance of the last four residues for the normal function of CFTR is also demonstrated by the occurrence of a stop mutation at Gln1476 in a patient with cystic fibrosis [http://genet.sickkids.on.ca]. Nevertheless,

additional sequences within the CFTR regions spanning residues 1370–1394 and 1404–1425 are also required for the apical localization of this channel [50]. Moreover, recent studies provided evidence that the PDZ-interacting sequence of CFTR is not an apical membrane-sorting motif but it controls the endocytic recycling and apical retention of CFTR [74]. Since several PDZ-containing proteins interact with this motif, it is difficult to dissect the contribution of each of these proteins in the endocytic recycling, apical localization/retention and activity of CFTR because of potential functional redundancy. For example, the lack of a phenotype associated with CFTR dysfunction in a targeted disruption of the *NHERF* gene in mice [69] can be attributed to functional compensation by NHERF2, CAP70 and/or other PDZ-containing proteins that interact with CFTR.

In addition to the apical membrane localization/retention, the bivalent binding of the NHERF PDZ domains to the CFTR C-terminal region was shown to activate the channel [61]. A similar effect on the CFTR activity was observed upon binding of the CAP70 region harboring the third and fourth PDZ domains of this protein to the CFTR C-terminal tail [80]. Since both PDZ domains of NHERF were required for the regulation of CFTR gating, it was proposed that these domains interact with distinct CFTR molecules, promoting channel dimerization and increasing the open probability of this channel [5, 61, 80]. Although the precise mechanism behind the channel activation brought about by dimerization is unknown, the demonstration that the NHERF and CAP70 binding to CFTR directly affected channel gating provided the first evidence that PDZ-mediated interactions may have regulatory functions, in addition to assembling transducisomes.

## Structural Determinants of the CFTR Interaction with the NHERF PDZ1 Domain

A first glimpse at the molecular recognition of CFTR by NHERF was provided by the crystal structure of the NHERF PDZ1 domain complexed with the CFTR C-terminal sequence QDTRL determined at 1.7 Å resolution [39]. The overall topology of NHERF PDZ1 is similar to other PDZ structures, consisting of six β-strands (β1–β6) and two α-helices (α1 and α2) (Fig. 1). The strands comprise an anti-parallel β-sandwich with one β-sheet formed by β1, β6, β4, and β5, and the second β-sheet formed by β2, β3, and β4 strands. The fold is stabilized by hydrophobic interactions involving the conserved residues Leu17, Phe26, Leu28, Ile39, Leu53, Leu59, Val76, Ile79, Val86, and Leu88, which form the core of the molecule. The CFTR peptide inserts into the PDZ1 binding pocket antiparallel to the β2 strand
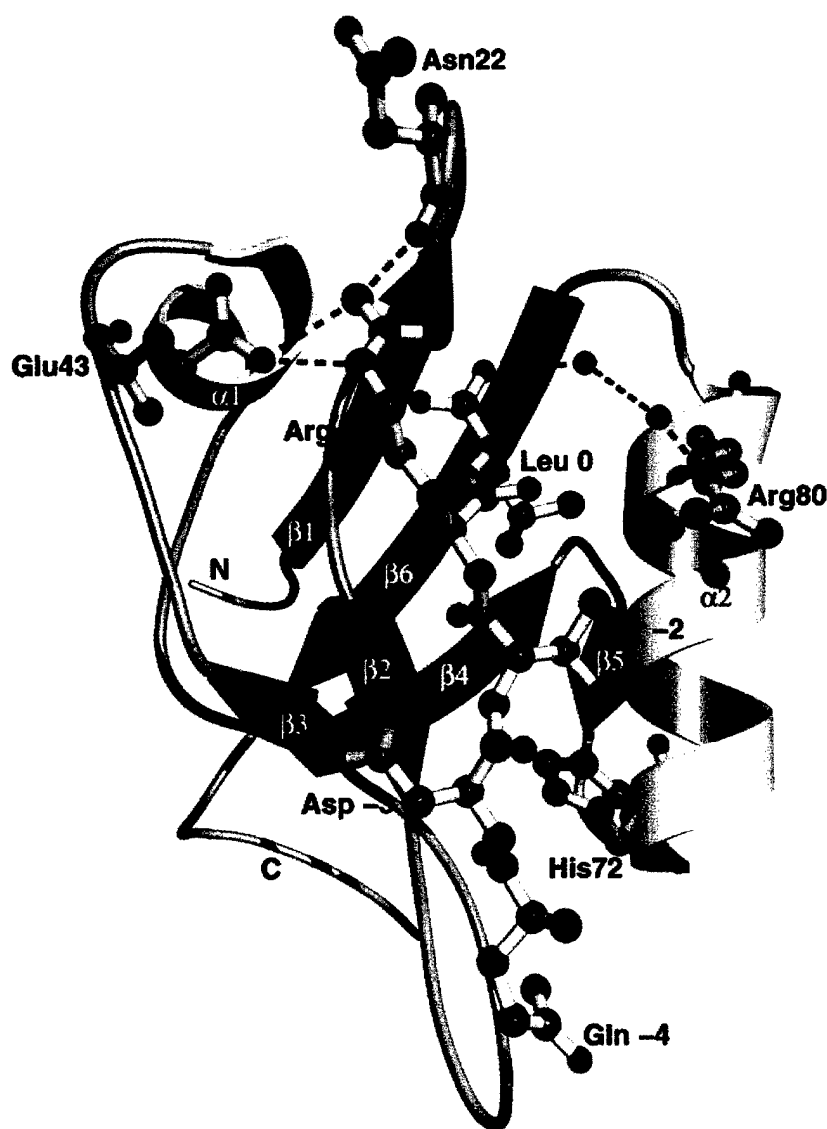
**Fig. 1.** Ribbon diagram of the human NHERF PDZ1 domain bound to the CFTR C-terminal sequence QDTRL. The β-strands are colored *brown* and the α-helices *yellow*. The peptide ligand is shown as a *white* ball-and-stick model. The PDZ residues Asn22, Glu43, His72, and Arg80 that participate in hydrogen bonding with peptidic side chains are shown as *pink* ball-and-stick models. Carbon, oxygen, and nitrogen atoms are shown in *black*, *red*, and *blue*, respectively. Water molecules are shown as *green spheres* and hydrogen bonds as *cyan dashed lines*. The figure was made using the published atomic coordinates (Protein Data Bank code 1i92).

and extends the β-sheet of PDZ1. In this arrangement, the invading pentapeptide is highly ordered, as indicated by low temperature factors. The side chain of the peptidic Gln −4 does not make contacts with PDZ1 residues and only the carbonyl oxygen forms a hydrogen bond with the amide nitrogen of Gly30, indicating that Gln −4 does not contribute to the specificity of the interaction. By contrast, Asp −3, Thr −2, and Leu 0 are engaged in numerous interactions with PDZ1, consistent with biochemical evidence on the important roles of these residues in the specificity and affinity of the NHERF PDZ1-CFTR interaction. Specifically, the side chain of Asp −3 forms a hydrogen bond with His27 and a salt bridge with Arg40. Similarly, the amide nitrogen and carbonyl oxygen of Thr −2 hydrogen bond with the carbonyl oxygen and amide nitrogen of Leu28, respectively, while the side chain of Thr −2 hydrogen bonds with the imidazole group of His72 (Fig. 1). The latter interaction corroborates the critical role of a threonine or serine residue at position −2 of the ligand and a conserved histidine

at the beginning of the α2 helix for the specificity of class I PDZ-peptide interactions [26, 28, 35, 68, 73, 79].

The side chain and carboxylate group of Leu 0 enter into a hydrophobic pocket formed by the NHERF residues Tyr24, Gly25, Phe26, Leu28, Val76, and Ile79. The isobutyl group of Leu 0 makes hydrophobic contacts with Phe26 and Ile79. In addition, the carboxyl oxygen of Leu 0 is engaged in hydrogen bonding with the amide nitrogen atoms of Gly25 and Phe26, whereas the carbonyl oxygen of Leu 0 hydrogen bonds directly with the amide nitrogen of Tyr24 and indirectly with the guanido group of Arg80 in the α2 helix through two ordered water molecules (Fig. 1). The involvement of Arg80 in carboxylate binding through ordered water molecules differs from other PDZ structures where this function is mediated by an arginine residue in the β1-β2 loop [15, 28], corresponding to NHERF PDZ1 Lys19. Interestingly, the NHERF Lys19 does not participate in hydrogen bonding with the terminal carboxylate group, indicating that there are signifi-

cant variations in the atomic structural determinants of the PDZ-ligand interactions.

Previous biochemical studies demonstrated that substitution of the C-terminal leucine with valine in peptide ligands markedly reduced binding to the NHERF PDZ1 domain [26, 27, 79]. Likewise, replacement of the highly conserved Leu1480 in CFTR by alanine abrogated the apical localization of this channel due to abolishment of the CFTR-PDZ interaction [53]. How does the NHERF PDZ1 domain discriminate between the side chain of a C-terminal leucine and smaller side chains like those of valine or alanine? The NHERF PDZ1-CFTR structure showed that the isobutyl group of Leu 0 makes several hydrophobic contacts in the PDZ1 carboxylate-binding pocket [39], suggesting that the hydrophobic character of this cavity would likely exclude polar and charged side chains. Furthermore, comparison of the NHERF PDZ1 and PSD-95 PDZ3 crystal structures complexed with peptide ligands having C-terminal leucine and valine residues, respectively, revealed that their hydrophobic pockets have different sizes and shapes [38]. The NHERF PDZ1 cavity is large and the isobutyl group of Leu 0 fits snugly in this pocket, whereas the smaller side chains of valine and alanine would leave vacated spaces within this cavity that would be energetically unfavorable [17]. Thus, it appears that the tight fit of the leucine side chain in the hydrophobic cavity provides an explanation for the strict requirement for C-terminal leucine in all the high-affinity ligands of NHERF PDZ1, and the poor affinity of this domain for C-terminal valine and alanine residues. By contrast, the smaller cavity of PSD-95 PDZ3 interacts tightly with the isopropyl group of valine, making the accommodation of the larger isobutyl group of leucine stereochemically challenging. Therefore, the sequence variation among different PDZ domains generates hydrophobic cavities with distinct volumes and shapes, providing a selectivity mechanism for ligand recognition based on the stereochemical complementarity of the peptidic C-terminal residue and the volume/shape of the cavity. In this context, it is important to note that both NHERF PDZ1 and PSD-95 PDZ3 are currently considered as class I domains despite their fundamental differences in discriminating between the C-terminal residues of their cognate ligands. This underscores the problem with the current classification of PDZ domains and provides a compelling argument for a more elaborate reclassification scheme that would take into account the exquisite ligand selectivity of these modules.

## The Importance of Arg −1 for the NHERF PDZ1-CFTR Interaction

The NHERF PDZ1-CFTR crystal structure also revealed a novel multivalent interaction of the arginine at position −1 of the CFTR peptide with two PDZ1 residues [39]. As mentioned above, early PDZ-peptide selection studies concluded that the residue at position −1 of the peptide ligand makes no contribution to the specificity and affinity of the PDZ-peptide interaction. Nevertheless, subsequent biochemical studies demonstrated that arginine is the preferred residue at position −1 for optimal binding to NHERF PDZ1. For example, affinity selection experiments showed that NHERF PDZ1 selected almost exclusively ligands with arginine at position −1 [79]. In addition, point mutagenesis of the penultimate arginine to alanine, phenylalanine, leucine, or glutamic acid, decreased the affinity of the PDZ1-ligand interaction [26]. However, the structural basis for the NHERF PDZ1 ability to discriminate between different side chains at the −1 position of the peptide remained obscure until the structural analysis of the NHERF PDZ1-CFTR complex. This crystal structure revealed that the guanido group of Arg −1 forms two salt bridges with the side chain of Glu43 and two hydrogen bonds with the carbonyl oxygen of Asn22 (Fig. 1). These interactions provided the first structural explanation for the remarkable preference for a penultimate arginine by NHERF PDZ1 and consolidated previous biochemical results on the importance of this amino acid in the affinity of the interaction. Importantly, involvement of the penultimate residue in the PDZ-ligand interaction is not exclusive to the NHERF-CFTR complex but it seems to represent a more general theme in the selectivity mechanisms of other PDZ domains. For example, the PDZ2 domain of the membrane-associated guanylate kinase MAGI3 also binds preferentially to ligands having a tryptophan at position −1 [22], and the PDZ1 domain of the scaffolding protein INAD forms a disulfide bond with the penultimate cysteine of the peptide ligand [41]. Therefore, it appears that PDZ domains have a preference for specific side chains at position −1 and interact optimally with peptide ligands having the corresponding penultimate residues. The NHERF-CFTR structure also allows the prediction that the penultimate arginine of other ligands that interact with NHERF PDZ1, such as the C-terminal sequence TRL of the Na/P$_i$-cotransporter IIa [25, 29], is involved in similar networks of salt bridges and hydrogen bonds with the Glu43 and Asn22 residues of NHERF.

## Similarities and Differences between the NHERF PDZ1 Interaction with CFTR and Membrane Receptors

Two recent crystal structures of the NHERF PDZ1 domain complexed with the C-terminal regions of β$_2$AR and PDGFR provided new structural insights into the contribution of the penultimate peptidic res-

idue to the affinity of the PDZ-ligand interaction [40]. In these structures the isobutyl group of Leu −1 and the phenolic ring of Phe −1 of the $\beta_2$AR and PDGFR ligands, respectively, engage in hydrophobic interactions with several PDZ1 residues. The side chains of Phe −1 and Leu −1 follow a path similar to that of the aliphatic portion of the Arg −1 side chain in the CFTR-PDZ1 structure, facing towards the PDZ residues Asn22 and Glu43. These two PDZ residues exhibit large conformational changes and they seem to play a critical role in the ability of NHERF PDZ1 to accommodate ligands with penultimate side chains of different hydrophobicity and polarity [40]. It remains to be seen whether the corresponding residues of other PDZ domains have similar roles in ligand recognition.

The three crystal structures of NHERF PDZ1 bound to CFTR, $\beta_2$AR, and PDGFR C-terminal tails represent the first structural analysis of a PDZ domain bound to three different ligands and provide an opportunity to identify significant differences in the PDZ-ligand interactions [39, 40]. One important difference was observed in the PDZ1-$\beta_2$AR structure, where the side chain of Asn −4 makes two hydrogen bonds with the amide nitrogen and carbonyl oxygen of Gly30, respectively, that contribute to the affinity of this interaction. By contrast, the side chains of the residues at position −4 of the CFTR and PDGFR ligands do not interact with PDZ1 amino acids. Another difference among these three structures is that the carbonyl oxygen atoms of the penultimate residues of both $\beta_2$AR and PDGFR ligands make direct hydrogen bonds with the guanido group of Arg80, whereas in the PDZ1-CFTR complex the carbonyl oxygen of Arg −1 does not hydrogen bond with Arg80. In addition, the carbonyl oxygen of Leu 0 interacts indirectly with Arg80 through two ordered water molecules in the PDZ1-$\beta_2$AR and PDZ1-CFTR but not in the PDZ1-PDGFR complex. Thus, the structural analysis of NHERF PDZ1 bound to CFTR, $\beta_2$AR, and PDGFR C-terminal sequences demonstrated that the ordered water molecules and hydrogen bond networks stabilizing the PDZ-ligand interaction differ even for slightly different ligands bound to the same PDZ domain.

**Future Directions**

The structural analysis of CFTR is in its infancy. Although remarkable progress has been made in our understanding of CFTR function during the past decade, structural studies of this channel and its complexes with regulatory proteins will undoubtedly revolutionize the field. Because crystallization of the full-length CFTR is a daunting task, crystallographic analyses of individual domains and domain complexes with regulatory and scaffolding proteins, such as syntaxin 1A, NHERF, NHERF2, and CAP70, is

an alternative approach that will yield extremely useful information about the function and regulation of this channel. Importantly, high-resolution atomic models of CFTR domains will be instrumental in determining the structures of full-length CFTR crystals, when they become available.

The crystal structure of the NHERF PDZ1 domain bound to the C-terminal region of CFTR has provided important insights into the molecular determinants of an interaction interface that is critical for the apical localization/retention and gating of the CFTR channel. At the same time however, it underscores the urgent need for more structural studies of protein complexes involving larger CFTR and NHERF fragments. Arguably, among the most pressing questions to be addressed is the structural basis of the mechanism behind the regulation of CFTR gating by the two PDZ domains of NHERF [61]. At present, the spatial arrangement of the NHERF PDZ domains in relation to each other and the mode of their interaction with two distinct CFTR molecules to induce channel dimerization are unknown. Furthermore, the somewhat controversial issue of the CFTR C-terminal recognition by the NHERF PDZ2 domain also needs to be resolved structurally. While early studies suggested that NHERF PDZ2 has a selectivity for sequence motifs different than that present in the CFTR C-terminal tail [79], a protein fragment spanning residues 132–299 of NHERF and encompassing the PDZ2 domain was shown to interact with the CFTR CTD [61]. The latter finding raises the intriguing possibility that sequences outside the NHERF PDZ2 borders may participate in interactions with CFTR residues located upstream of the C-terminal four amino acids. Therefore, structures of NHERF protein fragments containing both PDZ1 and PDZ2 domains complexed with the entire CFTR CTD are required to elucidate the molecular mechanism of CFTR gating by NHERF. Of course, similar structural studies of the CFTR CTD bound to other PDZ-containing proteins, including NHERF2 and CAP70, will provide equally important mechanistic insights into the CFTR regulation.

In addition to promoting dimerization of CFTR molecules, the multi-PDZ scaffolding proteins NHERF, NHERF2, and CAP70 may also link this chloride channel to a wide variety of transporters, ion channels, kinases, phosphatases, and cytoskeletal elements. It is well established that CFTR regulates the activity of a growing list of transporters and ion channels, including $Na^+/H^+$ exchangers, $Cl^-/HCO_3^-$ exchangers, epithelial $Na^+$ channels, renal $K^+$ channels, outwardly rectifying $Cl^-$ channels, and $Ca^{+-}$ activated $Cl^-$ channels (reviewed in [44, 67]). Although it is not known how CFTR regulates the activities of so many ion channels and transporters, an attractive possibility emanating from the multi-

plicity of PDZ-containing proteins that associate with CFTR is that multi-PDZ proteins orchestrate combinatorial interactions of CFTR with other channels to modulate their function. In support of this scenario, the regulatory interaction between CFTR and $Na^+/H^+$ exchanger 3 requires the C-terminal PDZ-binding motif of CFTR, suggesting that these proteins are organized into supramolecular complexes by NHERF or other PDZ-containing proteins [1]. Furthermore, the recent identification of the chloride channel ClC-3B as a target of NHERF PDZ2 [60] supports the hypothesis that NHERF may organize the assembly of a ternary complex containing CFTR and ClC-3B through its two PDZ domains. Interestingly, coexpression of CFTR with NHERF and ClC-3B in epithelial cells resulted in ClC-3B-dependent outwardly rectifying chloride channel activity regulated by CFTR [60], providing evidence for a critical role of NHERF in this interchannel regulation. Similar macromolecular complexes assembled by multi-PDZ proteins have been described in other systems, as exemplified by the coupling of the ionic channels TRP and TRPL to multiple signaling molecules by the five PDZ-containing scaffolding protein INAD in *Drosophila* retinal cells [19, 28, 68, 76]. Therefore, defining the structural basis of the physical interactions among CFTR, NHERF, and other proteins within these multi-component complexes will reveal their molecular relationships and elucidate the mechanisms underlying CFTR function at the atomic level.

From a clinical perspective, structural information on the CFTR channel and its complexes with regulatory proteins may have important implications for the development of molecular medical approaches for treating CFTR-associated diseases. For cystic fibrosis it would be desirable to develop CFTR agonists that would enhance the activity of mutant CFTR proteins, in particular the CFTR-$\Delta$F508, which harbors a deletion of Phe508 and represents the most common mutation in cystic fibrosis accounting for about 70% of all disease-causing alleles [6]. The CFTR-$\Delta$F508 protein does not fold correctly and is retained in the endoplasmic reticulum where it is targeted for degradation, leading to low channel density and reduced chloride transport in the apical membrane of epithelial cells [2]. Because CFTR-$\Delta$F508 can function as a chloride channel when expressed in the plasma membrane [16, 47, 65], it is conceivable, at least in theory, that development of CFTR agonists acting through NHERF to promote apical localization/retention and dimerization of CFTR-$\Delta$F508 might augment channel activity in these patients. A similar approach could be also used in other cystic fibrosis-associated mutations that affect the membrane localization of the CFTR channel. In practice however, it may prove difficult to develop agonists that would increase the affinity of NHERF

for CFTR because the sequence DTRL is considered to be the optimum ligand for NHERF PDZ1 [26]. Nonetheless, since it is possible that additional residues in the CFTR CTD also interact with the NHERF PDZ1-PDZ2 region, it could be envisioned that these interaction interfaces might provide more amenable targets for developing strategies to increase the affinity of the NHERF-CFTR interaction. Importantly, recent studies have identified a PDZ-containing Golgi-associated protein designated CAL [11], also known as PIST [59] or FIG [9], which modulates the membrane expression of CFTR. The CAL PDZ binds to the C-terminal tail of CFTR and promotes retention of this channel within the cell [11], suggesting that inhibitors of the CAL-CFTR interaction may increase the CFTR traffic to the apical membrane. Therefore, structural studies of the CFTR CTD bound to the PDZ domains of NHERF and other CFTR-associated proteins such as CAL, will reveal the three-dimensional interaction interfaces of these proteins and will provide structural frameworks for developing novel approaches aiming at enhancing CFTR activity. Likewise, structural analysis of the CFTR NTD/syntaxin 1A/SNAP-23 complex would guide the design of small-molecule compounds that block this interaction and enhance the chloride channel activity in cystic fibrosis patients carrying partial-loss-of-function mutations in the *CFTR* gene.

Structural studies of CFTR are also crucial for the development of new treatments for secretory diarrhea. Bacterial toxins that induce cyclic nucleotide production in the intestine promote CFTR phosphorylation and channel hyperactivation, which in turn results in massive secretion of salt and water [2, 20]. For example, activation of the PKA and cGMP-dependent protein kinase II by the cholera toxin and heat-stable enterotoxin from *E. coli*, respectively, results in overstimulation of CFTR, which subsequently leads to intestinal fluid and electrolyte secretion (secretory diarrhea) and dehydration. The role of NHERF and other PDZ-containing proteins in the apical membrane localization/retention and function of CFTR could be exploited for the development of novel CFTR inhibitors, which would act by blocking these interactions. Towards this goal, the atomic structures of the NHERF PDZ domains complexed with the CFTR CTD would guide the structure-based design of CFTR-PDZ blockers, and hopefully lead to the development of new therapeutics against secretory diarrhea.

A recent flurry of publications describing the crystallographic analysis of ion channels and ABC transporters leaves no doubt that we are entering a new era of high-resolution structural characterization of these membrane proteins. The unique properties of CFTR, being a chloride channel, a regulator of other channels, and a member of the ABC superfamily, in combination with its central role in the pathogenesis

of serious human diseases, make it a highly privileged target for structural studies. It is hoped that the wealth of structural and functional information to be discovered in the coming years on the regulation of CFTR trafficking, apical localization, and gating will lead to the development of novel ways to modulate channel function that may have clinical applications in treating CFTR-associated diseases.

## References

1. Ann, W., Kim, K.H., Lee, J.A., Kim, J.Y., Choi, J.Y., Moe, O.W., Milgram, S.L., Muallem, S., Lee, M.G. 2001. *J. Biol. Chem.* **276**:17236–17243

2. Akabas, M.H. 2000. *J. Biol. Chem.* **275**:3729–3732

3. Aleksandrov, L., Aleksandrov, A.A., Chang, X.B., Riordan, J.R. 2002. *J. Biol. Chem.* **277**:15419–15425

4. Bezprozvanny, I., Maximov, A. 2001. *FEBS Lett.* **509**:457–462

5. Bezprozvanny, I., Maximov, A. 2001. *Proc. Natl. Acad. Sci. USA* **98**:787–789

6. Bobadilla, J.L., Macek, M., Fine, J.P., Farrell, P.M. 2002. *Hum. Mutat.* **19**:575–606

7. Borst, P., Elferink, R.O. 2002. *Annu. Rev. Biochem.* **71**:537–592

8. Chang, G., Roth, C.B. 2001. *Science* **293**:1793–1800

9. Charest, A., Lane, K., McMahon, K., Housman, D.E. 2001. *J. Biol. Chem.* **276**:29456–29465

10. Chen, J.H., Chang, X.B., Aleksandrov, A.A., Riordan, J.R. 2002. *J. Membrane Biol.* **188**:55–71

11. Cheng, J., Moyer, B.D., Milewski, M., Loffing, J., Ikeda, M., Mickle, J.E., Cutting, G.R., Li, M., Stanton, B.A., Guggino, W.B. 2002. *J. Biol. Chem.* **277**:3520–3529

12. Cormet-Boyaka, E., Di, A., Chang, S.Y., Naren, A.P., Tousson, A., Nelson, D.J., Kirk, K.L. 2002. *Proc. Natl. Acad. Sci. USA* **99**:12477–12482

13. Dean, M., Hamon, Y., Chimini, G. 2001. *J. Lipid Res.* **42**:1007–1017

14. Diederichs, K., Diez, J., Greller, G., Muller, C., Breed, J., Schnell, C., Vonrhein, C., Boos, W., Welte, W. 2000. *EMBO J.* **19**:5951–5961

15. Doyle, D.A., Lee, A., Lewis, J, Kim, E., Sheng, M., MacKinnon, R. 1996. *Cell* **85**:1067–1076

16. Egan, M.E., Glockner-Pagel, J., Ambrose, C., Cahill, P.A., Pappoe, L., Balamuth, N., Cho, E., Canny, S., Wagner, C.A., Geibel, J., Caplan, M.J. 2002. *Nat. Med.* **8**:485–492

17. Eriksson, A.E., Baase, W.A., Zhang, X.J., Heinz, D.W., Blaber, M., Baldwin, E.P., Matthews, B.W. 1992. *Science* **255**:178–183

18. Eskandari, S., Wright, E.M., Kreman, M., Starace, D.M., Zampighi, G.A. 1998. *Proc. Natl. Acad. Sci. USA* **95**:11235–11240

19. Fanning, A.S., Anderson, J.M. 1999. *J. Clin. Invest.* **103**:767–772

20. Field, M., Semrad, C.E. 1993. *Annu. Rev. Physiol.* **55**:631–655

21. Fu, J., Ji, H.L., Naren, A.P., Kirk, K.L. 2001. *J. Physiol.* **536**:459–470

22. Fuh, G., Pisabarro, M.T., Li, Y., Quan, C., Lasky, L.A., Sidhu, S.S. 2000. *J. Biol. Chem.* **275**:21486–21491

23. Gadsby, D.C., Nairn, A.C. 1999. *Physiol. Rev.* **79**:S77–S107

24. Gaudet, R., Wiley, D.C. 2001. *EMBO J.* **20**:4964–4972

25. Gisler, S.M., Stagljar, I., Traebert, M., Bacic, D., Biber, J., Murer, H. 2001. *J. Biol. Chem.* **276**:9206–9213

26. Hall, R.A., Ostedgaard, L.S., Premont, R.T., Blitzer, J.T., Rahman, N., Welsh, M.J., Lefkowitz, R.J. 1998. *Proc. Natl. Acad. Sci. USA* **95**:8496–8501

27. Hall, R.A., Premont, R.T., Chow, C.W., Blitzer, J.T., Pitcher, J.A., Claing, A., Stoffel, R.H., Barak, L.S., Shenolikar, S., Weinman, E.J., Grinstein, S., Lefkowitz, R.J. 1998. *Nature* **392**:626–630

28. Harris, B.Z., Lim, W.A. 2001. *J. Cell Sci.* **114**:3219–3231

29. Hernando, N., Deliot, N., Gisler, S.M., Lederer, E., Weinman, E.J., Biber, J., Murer, H. 2002. *Proc. Natl. Acad. Sci. USA* **99**:11957–11962

30. Higgins, C.F. 1992. *Annu. Rev. Cell. Biol.* **8**:67–113

31. Higgins, C.F., Linton, K.J. 2001. *Science* **293**:1782–1784

32. Hillier, B.J., Christopherson, K.S., Prehoda, K.E., Bredt, D.S., Lim, W.A. 1999. *Science* **284**:812–815

33. Holland, I.B., Blight, M.A. 1999. *J. Mol. Biol.* **293**:381–399

34. Hopfner, K.P., Karcher, A., Shin, D.S., Craig, L., Arthur, L.M., Carney, J.P., Tainer, J.A. 2000. *Cell* **101**:789–800

35. Hung, A.Y., Sheng, M. 2002. *J. Biol. Chem.* **277**:5699–5702

36. Hung, L.W., Wang, I.X., Nikaido, K., Liu, P.Q., Ames, G.F., Kim, S.H. 1998. *Nature* **396**:703–707

37. Karpowich, N., Martsinkevich, O., Millen, L., Yuan, Y.R., Dai, P.L., MacVey, K., Thomas, P.J., Hunt, J.F. 2001. *Structure* **9**:571–586

38. Karthikeyan, S., Leung, T., Birrane, G., Webster, G., Ladias, J.A.A. 2001. *J. Mol. Biol.* **308**:963–973

39. Karthikeyan, S., Leung, T., Ladias, J.A.A. 2001. *J. Biol. Chem.* **276**:19683–19686

40. Karthikeyan, S., Leung, T., Ladias, J.A.A. 2002. *J. Biol. Chem.* **277**:18973–18978

41. Kimple, M.E., Siderovski, D.P., Sondek, J. 2001. *EMBO J.* **20**:4414–4422

42. Kirk, K.L. 2000. *Cell. Mol. Life. Sci.* **57**:623–634

43. Kocher, O., Comella, N., Tognazzi, K., Brown, L.F. 1998. *Lab. Invest.* **78**:117–125

44. Kunzelmann, K., Schreiber, R. 1999. *J. Membrane Biol.* **168**:1–8

45. Locher, K.P., Lee, A.T., Rees, D.C. 2002. *Science* **296**:1091–1098

46. Luo, J., Pato, M.D., Riordan, J.R., Hanrahan, J.W. 1998. *Am. J. Physiol.* **274**:C1397–C1410

47. Maitra, R., Shaw, C.M., Stanton, B.A., Hamilton, J.W. 2001. *Am. J. Physiol.* **280**:C1031–C1037

48. Marshall, J., Fang, S., Ostedgaard, L.S., O'Riordan, C.R., Ferrara, D., Amara, J.F., Hoppe, H., Scheule, R.K., Welsh, M.J., Smith, A.E., Cheng, S.H. 1994. *J. Biol. Chem.* **269**:2987–2995

49. Maudsley, S., Zamah, A.M., Rahman, N., Blitzer, J.T., Luttrell, L.M., Lefkowitz, R.J., Hall, R.A. 2000. *Mol. Cell. Biol.* **20**:8352–8363

50. Milewski, M.I., Mickle, J.E., Forrest, J.K., Stafford, D.M., Moyer, B.D., Cheng, J., Guggino, W.B., Stanton, B.A., Cutting, G.R. 2001. *J. Cell Sci.* **114**:719–726

51. Mohler, P.J., Kreda, S.M., Boucher, R.C., Sudol, M., Stutts, M.J., Milgram, S.L. 1999. *J. Cell Biol.* **147**:879–890

52. Moyer, B.D., Denton, J., Karlson, K.H., Reynolds, D., Wang, S., Mickle, J.E., Milewski, M., Cutting, G.R., Guggino, W.B., Li, M., Stanton, B.A. 1999. *J. Clin. Invest.* **104**:1353–1361

53. Moyer, B.D., Duhaime, M., Shaw, C., Denton, J., Reynolds, D., Karlson, K.H., Pfeiffer, J., Wang, S., Mickle, J.E., Milewski, M., Cutting, G.R., Guggino, W.B., Li, M., Stanton, B.A. 2000. *J. Biol. Chem.* **275**:27069–27074

54. Murthy, A., Gonzalez-Agosti, C., Cordero, E., Pinney, D., Candia, C., Solomon, F., Gusella, J., Ramesh, V. 1998. *J. Biol. Chem.* **273**:1273–1276

55. Naren, A.P., Cormet-Boyaka, E., Fu, J., Villain, M., Blalock, J.E., Quick, M.W., Kirk, K.L. 1999. *Science* **286**:544–548

56. Naren, A.P., Di, A., Cormet-Boyaka, E., Boyaka, P.N., McGhee, J.R., Zhou, W., Akagawa, K., Fujiwara, T., Thome, U., Engelhardt, J.F., Nelson, D.J., Kirk, K.L. 2000. *J. Clin. Invest.* **105**:377–386

57. Naren, A.P., Nelson, D.J., Xie, W., Jovov, B., Pevsner, J., Bennett, M.K., Benos, D.J., Quick, M.W., Kirk, K.L. 1997. *Nature* **390**:302–305

58. Naren, A.P., Quick, M.W., Collawn, J.F., Nelson, D.J., Kirk, K.L. 1998. *Proc. Natl. Acad. Sci. USA* **95**:10972–10977

59. Neudauer, C.L., Joberty, G., Macara, I.G. 2001. *Biochem. Biophys. Res. Commun.* **280**:541–547

60. Ogura, T., Furukawa, T., Toyozaki, T., Yamada, K., Zheng, Y.J., Katayama, Y., Nakaya, H., Inagaki, N. 2002. *FASEB J.* **16**:863–865

61. Raghuram, V., Mak, D.D., Foskett, J.K. 2001. *Proc. Natl. Acad. Sci. USA* **98**:1300–1305

62. Reczek, D., Berryman, M., Bretscher, A. 1997. *J. Cell Biol.* **139**:169–179

63. Riordan, J.R., Rommens, J.M., Kerem, B., Alon, N., Rozmahel, R., Grzelczak, Z., Zielenski, J., Lok, S., Plavsic, N., Chou, J.L., Drumm, M.L., Iannuzzi, M.C., Collins, F.S., Tsui, L.C. 1989. *Science* **245**:1066–1073

64. Rosenberg, M.F., Mao, Q., Holzenburg, A., Ford, R.C., Deeley, R.G., Cole, S.P. 2001. *J. Biol. Chem.* **276**:16076–16082

65. Rubenstein, R.C., Egan, M.E., Zeitlin, P.L. 1997. *J. Clin. Invest.* **100**:2457–2465

66. Schneider, E., Hunke, S. 1998. *FEMS Microbiol. Rev.* **22**:1–20

67. Schwiebert, E.M., Benos, D.J., Egan, M.E., Stutts, M.J., Guggino, W.B. 1999. *Physiol. Rev.* **79**:S145–S166

68. Sheng, M., Sala, C. 2001. *Annu. Rev. Neurosci.* **24**:1–29

69. Shenolikar, S., Voltz, J.W., Minkoff, C.M., Wade, J.B., Weinman, E.J. 2002. *Proc. Natl. Acad. Sci. USA* **99**:11470–11475

70. Shenolikar, S., Weinman, E.J. 2001. *Am. J. Physiol.* **280**:F389–F395

71. Sheppard, D.N., Welsh, M.J. 1999. *Physiol. Rev.* **79**:S23–S45

72. Short, D.B., Trotter, K.W., Reczek, D., Kreda, S.M., Bretscher, A., Boucher, R.C., Stutts, M.J., Milgram, S.L. 1998. *J. Biol. Chem.* **273**:19797–19801

73. Songyang, Z., Fanning, A.S., Fu, C., Xu, J., Marfatia, S.M., Chishti, A.H., Crompton, A., Chan, A.C., Anderson, J.M., Cantley, L.C. 1997. *Science* **275**:73–77

74. Swiatecka-Urban, A., Duhaime, M., Coutermarsh, B., Karlson, K.H., Collawn, J., Milewski, M., Cutting, G.R., Guggino, W.B., Langford, G., Stanton, B.A. 2002. *J. Biol. Chem.* **277**:40099–40105

75. Thomas, P.J., Hunt, J.F. 2001. *Nat. Struct. Biol.* **8**:920–923

76. Tsunoda, S., Sierralta, J., Sun, Y., Bodner, R., Suzuki, E., Becker, A., Socolich, M., Zuker, C.S. 1997. *Nature* **388**:243–249

77. Velarde, G., Ford, R.C., Rosenberg, M.F., Powis, S.J. 2001. *J. Biol. Chem.* **276**:46054–46063

78. Voltz, J.W., Weinman, E.J., Shenolikar, S. 2001. *Oncogene* **20**:6309–6314

79. Wang, S., Raab, R.W., Schatz, P.J., Guggino, W.B., Li, M. 1998. *FEBS Lett.* **427**:103–108

80. Wang, S., Yue, H., Derin, R.B., Guggino, W.B., Li, M. 2000. *Cell* **103**:169–179

81. Weinman, E.J., Steplock, D., Wang, Y., Shenolikar, S. 1995. *J. Clin. Invest.* **95**:2143–2149

82. Yuan, Y.R., Blecker, S., Martsinkevich, O., Millen, L., Thomas, P.J., Hunt, J.F. 2001. *J. Biol. Chem.* **276**:32313–32321

83. Zerhusen, B., Zhao, J., Xie, J., Davis, P.B., Ma, J. 1999. *J. Biol. Chem.* **274**:7627–7630